PB82-234188

The Sensitivity of Complex Photochemical Model Estimates to Detail in Information

Systems Applications, Inc.
San Rafael, CA

Prepared for

Environnmental Protection Agency
Research Triangle Park, NC

Sep 81

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-450/4-81-031a	2.	3. RECIPIENT'S ACCESSION NO. PB82 234188
4. TITLE AND SUBTITLE The Sensitivity of Complex Photochemical Model Estimates to Detail in Input Information		5. REPORT DATE 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) T. W. Tesche, C. Seigneur, L. E. Reid, P. M. Roth, W. R. Oliver, and J. C. Cassmassi		8. PERFORMING ORGANIZATION REPORT NO. SAI No. 330R-EF81-5
9. PERFORMING ORGANIZATION NAME AND ADDRESS Systems Applications, Incorporated 950 Northgate Drive San Rafael, California 94903		10. PROGRAM ELEMENT NO. 11. CONTRACT/GRANT NO. 68-02-2870
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency Office of Air Quality Planning and Standards Research Triangle Park, North Carolina 27711		13. TYPE OF REPORT AND PERIOD COVERED 14. SPONSORING AGENCY CODE
15. SUPPLEMENTARY NOTES		

16. ABSTRACT

Using the air quality, meteorological and emissions data base available in the Los Angeles area, two days with distinctly different meteorology are simulated using a photochemical grid model (Urban Airshed Model). The data base used to generate model inputs is then degraded for the purpose of noting which data are most essential to collect in order to have a complex grid model perform adequately. The results are used to develop a more general methodology for prioritizing data needs. The methodology considers model sensitivity to input derived from data bases of varying detail, expense in collecting the data, and the uncertainty associated with deriving model input variables from the data base.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Photochemical grid models Urban Airshed Model Ozone Sensitivity studies Model inputs		

18. DISTRIBUTION STATEMENT Unlimited	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES 186
	20. SECURITY CLASS *(This page)*	22. PRICE

PB82-234188

United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
Research Triangle Park NC 27711

EPA-450/4-81-031a
September 1981

Air

EPA

The Sensitivity Of Complex Photochemical Model Estimates To Detail In Input Information

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA 22161

This report was furnished to the U.S. Environmental Protection Agency by Systems Applications, Incorporated in fulfillment of Contract 68-02-2870. The contents of this report are reproduced as received from Systems Applications, Incorporated. The opinions, findings and conclusions expressed are those of the author and not necessarily those of the Environmental Protection Agency. Mention of company or product names is not to be considered as an endorsement by the Environmental Protection Agency.

EPA-450/4-81-031a

The Sensitivity Of Complex Photochemical Model Estimates To Detail In Input Information

EPA Project Officer: Edwin L. Meyer, Jr.

Prepared for

U.S. Environmental Protection Agency
Office of Air, Noise and Radiation
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina 27711

September 1981

PREFACE

This report and its companion appendixes present the results of a large number of air quality sensitivity simulations as well as supporting analyses. Given the volume of information contained herein, the reader may wish to first gain an overview of the study by reading the introduction and conclusions, chapters I and VII, respectively.

ACKNOWLEDGMENTS

Several individuals made significant contributions to the technical work presented in this report. Assistance from the technical staff of the California Air Resources Board--in particular, Messrs. Andrew Ranzieri, Paul Allen, and Bill Loscutoff--is gratefully acknowledged. Mr. Ron Taira of the California Department of Transportation was especially helpful in formulating and supplying inputs for certain motor vehicle emission inventories. Contributions from individuals at the Office of Air Quality Planning and Standards, notably Dr. Edwin Meyer, and Messrs. John Summerhayes, David Barrett, and James Southerland, were central to the direction of this study. Special appreciation is given to several members of the SAI technical staff, particularly Drs. G. Z. Whitten, S. D. Reynolds, and M. K. Liu, for their advice and consultation throughout the course of the study. Production of the final report was due largely to the dedicated efforts of Karyl Pullen and Carol Lawson, the technical editors, as well as several members of the Publications Center.

EXECUTIVE SUMMARY

Considerable effort has been devoted in recent years to the development of urban-scale photochemical simulation models for analyzing the impacts of alternative emission control strategies on ambient oxidant concentrations. Because of the complexity of the physical and chemical processes governing the formation, transport, and dispersion of secondary pollutants, complex grid-based atmospheric simulation models are attractive tools for relating air quality concentrations to pollutant emission levels. Although evidence related to the performance of these models is limited, it appears that uncertainties in the model predictions fall within the bounds of current accepted practice; computed ozone levels are, on the average, within 25 to 40 percent of the measured values when the measurements exceed the National Ambient Air Quality Standard of 0.12 ppm (12 pphm).

Clearly, reduction in the uncertainties associated with model performance is desirable; however, a main concern regarding the use of these urban-scale models for air quality simulation is the cost and complexity of this approach. To simulate urban photochemical pollution, a considerable amount of emission, meteorological, and air quality data is required, and the difficulties involved in acquiring such a data base have discouraged wider application of these models. It is quite important, therefore, to ascertain the extent to which the amount and quality of input data may be reduced from the level considered desirable in current practice, without substantially reducing the accuracy of the model predictions. To acquire such knowledge, studies of the sensitivity of urban-scale air quality models to the level of detail of the input data are required. Several sensitivity studies have been performed with the Systems Applications, Incorporated (SAI), Airshed Model and with the Livermore regional air quality model (LIRAQ). However, these studies focus on a limited number of sensitivity cases; they do not provide sufficient information about the relative importance of the various sets of input data (meteorology, air quality, emission inventories) on model performance.

In this report a general methodology is developed to define a priority ranking of the input data needed for urban airshed simulation, which is then applied to air quality simulations of the Los Angeles Basin.

METHODOLOGY FOR RANKING DATA NEEDS

The methodology presented here is based on the definition of a sensitivity-uncertainty index, and requires information on the sensitivity of the model to the level of input data detail (e.g., number of monitoring stations and emission inventory), on the uncertainty of the model input variables (e.g., wind direction and nitrogen oxide concentrations), and on the cost required for acquiring the input data. These factors are considered in the definition of a sensitivity-uncertainty index that provides the basis for ranking data needs. A sensitivity index may be defined as follows:

$$\Gamma_{ij} = \frac{\Delta J_{ij}}{\Delta I_j} \quad , \tag{1}$$

where ΔJ_{ij} is a measure of the deviation in model prediction, for instance, the absolute deviation in ozone levels for the input data j and base case simulation i, and ΔI_j is the corresponding perturbation in the input data.

Next, the sensitivity index is normalized with respect to a cost difference (ΔK_j) corresponding to the change in input data ΔI_j, because the ranking of data needs clearly must include such information. The normalized sensitivity index is then defined as follows:

$$S_{ij} = \Gamma_{ij} \frac{\Delta I_j}{\Delta K_j} = \frac{\Delta J_{ij}}{\Delta K_j} \quad . \tag{2}$$

This index is a measure of the effect of the cost (associated with improving the level of detail of input data j) on model predictions for episode i.

The need for increasing the detail of input data should be weighed, however, against the accuracy available for corresponding model input variables. Accordingly, the following uncertainty index is introduced:

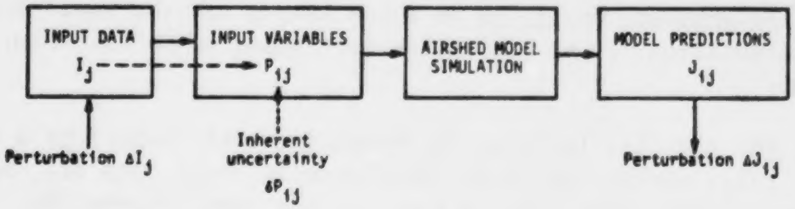
$$U_{ij} = \frac{\delta P_{ij}}{P_{ij}} \quad , \tag{3}$$

where δP_{ij} is the uncertainty in P_{ij} resulting from limitation in the representation of a mathematical or physico-chemical process. The model input variable P_{ij}, which has the greatest effect on model performance during the sensitivity study, is defined by i and j. For instance, in the case of a perturbation, ΔI_j is the number of meteorological stations and P_{ij} may refer to the wind field or the mixing height.

The larger the uncertainty in an input variable, the less additional input data is needed, since expenditure of considerable resources may not materially reduce the uncertainties associated with the input parameter. Hence, we weight the sensitivity index S_{ij}, by means of the uncertainty index U_{ij}, and a sensitivity-uncertainty index is defined as follows:

$$\Sigma_{ij} = \frac{S_{ij}}{U_{ij}} = \frac{\Delta J_{ij}}{\Delta K_{ij}} \frac{P_{ij}}{\delta P_{ij}} \quad . \tag{4}$$

A schematic representation of the sensitivity-uncertainty analysis is shown below.



SENSITIVITY STUDIES OF THE URBAN AIRSHED MODEL

The SAI Airshed Model has been applied to the urban areas of Los Angeles, Denver, Las Vegas, Sacramento, and St. Louis. Since the Los Angeles basin offered the most complete, readily available data base, it was chosen over other urban areas for the sensitivity studies. From the outset of the study, the influence of different meteorologies on the sensitivity study results was of concern. In Los Angeles, the close proximity of coastal and desert environs leads to shallow mixing layers (from 240 to 300 meters) along the coastal margin and to deeper mixing layers (from 900 to 1100 meters) farther inland during the midday hours. To determine whether the selection of different meteorological episodes might alter a subsequent ranking of data needs, two different air pollution episodes were considered: 26 June 1974 and 4 August 1975.

The level of airshed model performance for these two days is comparable to the general performance obtained in other air quality studies conducted with the model. The average absolute difference between predicted and observed concentrations for these two simulations is 31 to 35 percent (7.5 to 8.7 pphm) for ozone levels above 20 pphm.

A total of 22 sensitivity studies was carried out with these simulations of the Los Angeles basin air quality. The studies involved perturbations in the meteorological air quality and emission data, as well as simplification of the model structure. The resulting changes in the model predictions have been analyzed to provide information on model sensitivity.

Assessment of model sensitivity to input data may be considered in two parts:

- The sensitivity of the model output to changes in the input variables (e.g., wind fields, diffusivities, emission rates and mixing depth), may be evaluated through successive model simulations, each involving a prescribed set of input variables.
- The sensitivity of these input variables to perturbations in the basic data used to construct them (for instance, gridded wind field inputs to the SAI Airshed Model will vary, depending on the number of meteorological stations used to prepare wind model inputs.)

Thus, interpretation of airshed model sensitivity is seen as the combination of the sensitivity of the model to its input variables and of the sensitivity of the input preparation procedures (or models) to their basic data requirements. A complete discussion of the sensitivity results using this approach is presented in appendix A.

The sensitivity cases considered in this study are summarized in table 1. These simulations may be grouped into four general categories:

- Meteorological data: Sensitivity cases 1 through 4, and sensitivity case 5-2.
- Air quality data: Sensitivity cases 5 through 10.
- Emission data: Sensitivity cases 11 through 19.
- Model structure: Sensitivity cases 20 through 22.

An important step in quantifying model sensitivity is the definition of appropriate measures. The following measures are useful for this purpose:

Signed deviation is calculated as follows:

$$\frac{1}{N'}\sum_{j=1}^{N'}\frac{1}{N}\sum_{i=1}^{N}\frac{C_{s,i,j} - C_{b,i,j}}{C_{b,i,j}} \quad , \tag{5}$$

where $C_{s,i,j}$ and $C_{b,i,j}$ are the ozone concentrations for the sensitivity case and the base case, respectively, at station i for the hour j; N is the number of stations; and N' is the number of simulation hours.

Absolute deviation (in ozone levels) is calculated as follows:

$$\frac{1}{N'}\sum_{j=1}^{N'}\frac{1}{N}\sum_{i=1}^{N}\left|\frac{C_{s,i,j} - C_{b,i,j}}{C_{b,i,j}}\right| \quad . \tag{6}$$

These sensitivity measures are presented for ozone levels > 12 pphm in table 1. Other sensitivity measures were also considered: overall maximum ozone levels, the magnitude and time of occurrence of peak ozone levels at various stations, and dosage based on ozone levels above 12 pphm. Although results obtained with these different sensitivity measures are consistent, in some cases (e.g., for maximum ozone predictions) they are less pronounced.

Simulations involving a reduction in the upper air meteorological data or a change in the magnitude of reactive hydrocarbon initial conditions result in the greatest signed and absolute deviations. Such deviations range from 15 to 45 percent. The change in grid size results in an absolute deviation of about 20 percent. Other perturbations in the input data result in absolute deviations of less than 15 percent.

A reduction in meteorological data results in the lowest temporal correlations, with values as low as 0.023 and -0.183 for the 26 June 1974 simulations. Changes in the hydrocarbon initial conditions also result in a low temporal correlation coefficient for the 4 August 1975 simulation day. The same simulations (reduction in meteorological data, change in hydrocarbon initial conditions) also lead to low spatial correlation coefficients. The change in grid size results in a relatively low spatial correlation (i.e., $\rho_s < 0.721$), as expected, since reducing the grid

TABLE 1. SENSITIVITY MEASURES FOR OZONE CONCENTRATIONS ABOVE 12 PPHM

Sensitivity Study	Description*	Signed Deviation	Absolute Deviation
1	Reduced upper air meteorological data (J)	-0.177	0.300
2	Reduced upper air meteorological data (A)	-0.094	0.164
3	Reduced surface and upper air meteorological data (J)	-0.172	0.320
4	Reduced surface and upper air meteorological data (A)	-0.146	0.189
5-1	Reduced surface air quality data (J)	0.000	0.028
5-2	Reduced surface air quality, surface and upper air meteorological data (J)	-0.208	0.329
6	Reduced upper air quality data, more precursor material aloft (J)	0.105	0.106
7	Reduced upper air quality data, less precursor material aloft (A)	-0.073	0.073
8	Reduced surface air quality data (J)	-0.141	0.145
9	Reduced initial hydrocarbon concentration data (J)	-0.344	0.344
10	Reduced initial hydrocarbon concentration data (A)	-0.450	0.450
11	Reduced hydrocarbon emission speciation data (J)	0.046	0.046
12	Reduced hydrocarbon emission speciation data (A)	0.121	0.125
13	Outdated mobile source emissions (J)	0.027	0.037
14	Estimated mobile source emissions (J)	0.064	0.067
15	Estimated mobile source emissions (A)	0.115	0.127
16	Reduced data of point source emission temporal distribution (J)	-0.034	0.035
17	Reduced data of area source emission spatial distribution (J)	0.054	0.055
18	Reduced data of area source emission spatial distribution (A)	0.118	0.127
19	Reduced data of area source emission temporal distribution (J)	-0.005	0.007

Sensitivity Study	Description	Signed Deviation	Absolute Deviation
20	Larger grid size (J)	0.148	0.193
21	Two-grid layer model (instead of 4)	-0.007	0.053
22	One grid-layer model (instead of 4)	-0.067	0.207

---

* J and A refer to the 26 June 1974 and 4 August 1975 simulation, respectively.

resolution tends to spread out the predicted ozone concentration fields (i.e., the pollutant levels are more diffuse).

The large effect on model predictions of a reduction in upper air meteorological data is due in part to the fact that the wind field used with this urban air quality model is obtained by solving a boundary value problem for a potential function. Therefore, the simulation results of cases 1 and 2 are more sensitive to meteorological data located at the boundaries of the airshed than to the intensity of surface data acquisition in midbasin, for example.

Reductions in surface meteorological data (sensitivity cases 3 and 4), which affect mixing depths, wind direction, and wind speeds, lead to some perturbations in the ozone levels. However, the overall effect is not as large as that caused by reductions in the upper air meteorological data.

The reduction in the number of air quality monitoring stations (sensitivity case 5) does not have a major effect on model predictions. It should be noted, however, that the Los Angeles basin monitoring network is rather dense, and it is not surprising that the model is relatively insensitive to the intensity of surface air quality data. It appears, when comparing the sensitivity measures of cases 3, 5-1, and 5-2, that the reductions in air quality data and meteorological data are not additive. This reflects the complexity of chemical and physical processes in the atmosphere and the importance of including in a mathematical model all relevant phenomena.

An increase and reduction in precursor levels aloft lead to higher and lower ozone levels, respectively, as expected (sensitivity cases 6 and 7). Similarly, lower ozone levels result when the initial and boundary conditions are reduced to background levels (sensitivity case 8).

The deviations in ozone levels due to the prescription of initial non-methane hydrocarbon concentrations are of the order of 30 to 40 percent. This large perturbation confirms the idea that these data are an important input to an urban photochemical model.

In sensitivity case 11, an average hydrocarbon speciation (representative of motor vehicle emissions) was considered for all source categories. The main effect was to increase the chemical reactivity of the hydrocarbon emission mixture from petroleum refinery sources that, in turn, results in higher ozone levels downwind.

The sensitivity of the model to mobile source emissions has been investigated in cases 13, 14, and 15. In case 13, an outdated emission inventory was used; the perturbations in model predictions are relatively small. However, the use of a simpler inventory, based on fuel sales and demographic distributions, results in a greater effect on ozone levels predictions (up to 16 percent in absolute deviation for case 15), as demonstrated by the sensitivity measures of cases 14 and 15.

Changes in the temporal resolution of point source emissions (case 16) do not lead to any major perturbations in the simulations. The spatial resolution of area source emissions was modified in cases 17 and 18. Larger perturbations were observed with the 4 August 1975 conditions (case 18), and also for case 15 (in contrast to case 14) and case 12 (in contrast to case 11). This is due to the lower mixing heights on 4 August 1975 compared with 26 June 1974; consequently, a perturbation in the emission levels is associated with a smaller mixing volume on 4 August 1975, thus inducing larger deviations in pollutant concentrations. Changes in the temporal resolution of area source emissions lead to a negligible effect on model predictions, as indicated by the sensitivity measures of case 19.

The model structure was modified in the last three sensitivity cases. An enlargement of the grid size introduced an absolute deviation in ozone levels on the order of 20 percent. The reduction of the number of grid layers from 4 to 2 had relatively little effect on model performance. This simplification might be used in some model applications to reduce computational costs without degrading the accuracy of the simulations. It is important, however, to have one layer below the inversion and one layer above. Simulation with only one layer below the inversion lead to notable deviations in ozone levels.

It is also important to note at this point that the use of single-day simulations, in contrast to multiple-day simulations, distorts the importance of initial conditions and underestimates the importance of emissions. This distinction is interpreted in light of recent multiple-day airshed model simulations of the South Coast Air Basin (Killus et al. 1980; Souten et al. 1980).

APPLICATION OF THE METHODOLOGY TO THE LOS ANGELES BASIN

The methodology introduced earlier for ranking the input data needs of an urban airshed model was applied to the Los Angeles basin. From the results of the sensitivity studies presented previously, and from uncertaintly and cost estimates, sensitivity-uncertainty indexes were calculated. Sensitivity studies that involved the gridded structure of the model are not considered in this section.

The results of the sensitivity-uncertainty analyses are presented in table 2 for the 12 sets of input data, along with the corresponding ozone level deviations, estimated costs for data acquisition, and uncertainties in the model input variable.

Only the ranking of meteorological input data (upper air and surface) is affected by the choice of oxidant episode. Although the value of the sensitivity-uncertainty indexes of the other input data (air quality, chemistry, emissions) vary depending upon the episode, their relative ranking is not affected. This suggests that the meteorology of an urban area must be taken into account primarily when considering the need for meteorological input data.

Considering the 26 June 1974 sensitivity results, it is seen that updating the Los Angeles mobile source inventory ranks the highest, because the cost involved is relatively low and because the accuracy of mobile source emission rates is reasonable. Upper air data acquisition can be important; this depends, however, on the meteorology of the urban areas. For a single-day simulation, the initial conditions for reactive hydrocarbons are also ranked high, despite the relatively high costs of such a large monitoring effort. The latter, however, have been shown to be a key input to the airshed model. Predictions may be greatly affected by variations in hydrocarbon initial and boundary conditions. Reactive hydrocarbons, an essential precursor of photochemical smog, are difficult to measure accurately in ambient air; they constitute a small amount of the total mass of hydrocarbons, which is composed primarily of methane. They are usually estimated via assumed HC/NO_x ratios or as a specified fraction of the total hydrocarbon concentrations. Obviously, considerable uncertainty is involved in these empirical formulas.

Spatial resolution of area sources, upper air pollutant concentration data, detailed hydrocarbon speciation, and a detailed point source inventory generally are of similar importance to model performance. The importance of surface meteorology is of the same order, though it varies, depending on the episode simulated.

Surface air quality is ranked relatively low. This is because we assumed a 10-station network in the Los Angeles basin for analysis. The sensitivity relationship is probably nonlinear and, if only two or three stations had been considered, the importance of surface air quality would likely have been higher. This underscores the recommendation that the results of the sensitivity-uncertainty analysis must be considered in light of the assumptions made in the sensitivity study. It would be misleading to assume that surface air quality data have little effect on model predictions in other applications.

TABLE 2. SENSITIVITY-UNCERTAINTY INDEXES

Input Data	Ozone Level Deviation[*]	Cost for Data Acquisition (dollars)	Model Input Component Affected	Uncertainty	Sensitivity-Uncertainty Index[†] x 10^6
Upper air meteorology	0.33 (J) 0.17 (A)	50,000-100,000	Wind field (direction) Emissions source Mixing depth	25°-50° 10-50% 50-70%	6.6 - (21) - 66 2.4 - (4.1) - 6.8
Surface meteorology	0.007 (J) 0.06 (A)	20,000-30,000	Wind field Emissions sources	10-50%	0.5 - (1.3) - 3.5 4 - (11) - 30
Surface air quality	0.03 (J)	75,000-150,000	NO_x, RHC	20-60%	0.33 - (0.82) - 2
Upper air quality	0.11 (J)	50,000-125,000	NO_x, RHC	20-60%	1.5 - (4.1) - 11
IC for HC	0.32 (J) 0.43 (A)	75,000-150,000	RHC	20-60%	3.6 - (8.7) - 21.3 4.8 - (11.7) - 28.6
HC speciation	0.045 (J) 0.14 (A)	20,000-100,000	RHC	20-60%	0.75 - (2.9) - 11.25 2.3 - (9.0) - 35
Mobile source updating inventory	0.04 (J)	5,000-20,000	NO_x, RHC emissions.	10-30%	6.7 - (23) - 80
Mobile source gas sales	0.07 (J) 0.16 (A)	250,000-1,000,000	NO_x, RHC emissions.	10-30%	0.23 - (0.8) - 2.8 0.53 - (1.8) - 6.4
Point sources	0.01 (J)	10,000-50,000	NO_x, RHC emissions.	10-30%	0.67 - (2.6) - 10
Area sources spatial resolution	0.055 (J) 0.14 (A)	50,000-100,000	NO_x, RHC emissions.	10-30%	1.8 - (4.5) - 11 4.6 - (11.5) - 28
Area sources temporal resolution	.01 (J)	60,000-150,000	NO_x, RHC emissions.	10-30%	0.22 - (0.6) - 1.7

---

[*] J and A refer to the 26 June 1974 and 4 August 1975 simulations, respectively.

[†] Lower bound, geometric mean value, upper bound.

A detailed specification of transportation patterns is a rather expensive task, and this results in a very low ranking of the need for such input data. This result may be compared with the simple updating of the same inventory. This shows the importance of the assumptions that have been made and indicates that data needs will vary according to the existing data base. Quantification of the temporal distribution of stationary source emissions is expensive to obtain and appears to have little effect on model performance. This results in a low sensitivity-uncertainty index.

The sensitivity-uncertainty analysis that has been presented should be seen as a procedure for defining input data needs; the results obtained for the Los Angeles basin should be considered as an illustration of the method. It is recommended that costs and uncertainties be estimated on a site-specific basis for the photochemical model application area of interest; such cost and uncertainty estimates can then be used with the results of the sensitivity studies (ΔJs) to compute the sensitivity-uncertainty indexes.

The results of the sensitivity simulations depend on the perturbations introduced in the input data, the area chosen for the study, the atmospheric conditions of the prototype episode, and the modeling conditions (e.g., type of wind model and the duration of the simulation). The effect of atmospheric conditions has been considered for several sensitivity studies carried out under different meterological conditions (simulations of 26 June 1974 and 4 August 1975).

Limitations of the sensitivity studies result from the specific attributes of the Los Angeles basin and the length of simulation time. The influence of the air quality data and emission inventories on model predictions is affected by the length of simulation time. When the simulation is extended from a single-day to a multiple-day run, the importance of air quality data decreases, whereas the effect of changes in emission inventories on model predictions increases. It might be possible, however, to use the information obtained for the Los Angeles area in another location by evaluating the specific attributes of both areas. For instance, a classification of urban areas in terms of atmospheric attributes has recently been developed (Hillyer, 1980), and it should provide guidance in the generalization of these sensitivity results.

CONTENTS

PREFACE ii

ACKNOWLEDGMENTS iii

EXECUTIVE SUMMARY iv

LIST OF ILLUSTRATIONS xx

LIST OF TABLES xxi

I INTRODUCTION I-1

II CHARACTERISTICS OF THE SOUTH COAST AIR BASIN II-1

- A. Study Area Selection II-1
- B. Photochemical Oxidant Formation in the South Coast Air Basin II-2
  - 1. Meteorology: The Inversion and the Sea Breeze II-3
  - 2. Emission Distribution and Oxidant Formation II-4
  - 3. Effects of Topography on Pollutant Transport II-5
  - 4. The SCAB as a Case Study II-6
- C. Description of Two Los Angeles Oxidant Episodes II-7
- D. Available Data Resources II-7
  - 1. Emission Inventory II-12
  - 2. Aerometric Data II-12
- E. Previous Sensitivity Studies in the South Coast Air Basin II-13

III STRUCTURE OF THE SENSITIVITY SIMULATIONS III-1

- A. Generic Classification of Sensitivity Runs III-1
- B. Procedural Issues III-1
  - 1. Modeling Region III-1
  - 2. Multiple Base Cases III-5
  - 3. Simulation Duration III-5
  - 4. Selection of the Reduced Set of Air Quality Monitors III-7

C. Attributes of Each Sensitivity Run.................... III-9
  1. Simulations Focusing on a Limited Number of Aerometric Monitoring Stations.................... III-9
  2. Simulations Focusing on More Specialized Aerometric Monitoring Activities.................. III-10
  3. Simulations Focusing on Details in Emissions Inventories.................................... III-12
  4. Simulations Focusing on Model Grid Mesh Configuration........................... III-30
D. Concluding Remarks................................ III-31

IV BASE-CASE SIMULATIONS.............................. IV-1
A. Identification of Analysis Procedures................ IV-1
B. Base-Case Simulation Results......................... IV-4
  1. Accuracy of Computed Peak Concentrations........... IV-4
  2. Estimates of Systematic Bias....................... IV-7
  3. Estimates of Gross Error........................... IV-11
  4. Temporal Correlation............................... IV-14
  5. Spatial Alignment.................................. IV-15
C. Simulation Results for Specific Monitoring Stations.. IV-15
D. Evaluation of Ground-Level Ozone Concentration Fields.. IV-30
E. Summary.. IV-45

V SIMULATION RESULTS.................................. V-1
A. Introduction... V-1
  1. Review of Sensitivity Analyses of Air Quality Models................................ V-1
  2. Sensitivity of the Airshed Model to Input Data..... V-2
B. Measures for Ascertaining Model Sensitivity........... V-2
  1. Signed Deviation................................... V-4
  2. Absolute Deviation................................. V-5
  3. Temporal Correlation............................... V-5
  4. Spatial Correlation................................ V-6
  5. Overall Maximum Ozone Level........................ V-6
  6. Maximum Ozone Statistics........................... V-6
  7. Dosage... V-7
  8. Isopleths of Maximum Ozone Deviation............... V-7
  9. Ozone Profiles at Air Quality Monitoring Stations.. V-8
  10. Summary of Sensitivity Measures.................... V-8

C. Summary of Sensitivity Results........................ V-9
D. Sensitivity Results.................................. V-21

VI INTERPRETATION OF RESULTS............................... VI-1
A. Ranking of Data Needs through Sensitivity-Uncertainty Analysis............................ VI-1
1. Definition of a Sensitivity-Uncertainty Index...... VI-1
2. Cost Estimates for Data Acquisition............... VI-6
3. Uncertainty Estimates for Model Input Variables.... VI-6
4. Ranking of Data Needs............................. VI-9
5. Conclusions....................................... VI-18
B. Generalization of Results............................ VI-21
1. Specific Attributes of the Simulations............ VI-21
2. Limitations of the Sensitivity Results............ VI-24
3. Generalization of the Results..................... VI-28

VII CONCLUSIONS.. VII-1
A. General Procedure for Sensitivity Analysis........... VII-1
B. Specific Results of This Study....................... VII-3
C. Future Needs... VII-6

REFERENCES R-1

ILLUSTRATIONS

III-1 Modeling Region Used in the Airshed Model Sensitivity Study III-6

IV-1 Estimates of Systematic Bias in Computed Ozone Concentrations as a Function of Observed Concentration Level.. IV-10

IV-2 Estimates of Error in Computed Ozone Concentrations as a Function of Measured Concentration Levels.............. IV-12

IV-3 Distribution of Residuals (Predictions Minus Observations) for the Ozone Simulation Results.............. IV-13

IV-4 Number of Grid Cells for Which Model Predictions Bracket Observed Ozone Concentrations for 18 Stations during the Period 10:00 a.m. to 5:00 p.m. PST...................... IV-16

IV-5 Topographical Features and the Locations of Air Monitoring Stations in the South Coast Air Basin............ IV-19

IV-6 Calculated and Observed Ozone Concentrations for 26 June 1974.. IV-20

IV-7 Calculated and Observed Ozone Concentrations for 4 August 1975.. IV-25

IV-8 Ozone Isopleths for 26 June 1974............................ IV-31

IV-9 Ozone Isopleths for 4 August 1975........................... IV-38

V-1 Sensitivity Analysis of the Urban Airshed Model............. V-3

V-2 Signed Deviations for Ozone Concentrations Above 12 pphm... V-16

V-3 Absolute Deviations For Ozone Concentrations Above 12 pphm... V-17

VI-1 Comparison of "Perturbation Specific" and Classical Sensitivity Indexes for a Hypothetical Case.................. VI-4

VI-2 Sensitivity-Uncertainty Indexes--Simulation of 26 June 1974: Sensitivity to Ozone Levels above 12 pphm.. VI-14

VI-3 Sensitivity-Uncertainty Indexes--Simulation of 4 August 1975: Sensitivity to Ozone Levels above 12 pphm.. VI-15

VI-4 Sensitivity-Uncertainty Indexes--Simulation of 26 June 1974: Sensitivity to Air-Shed-Wide Peak Ozone Level.. VI-16

VI-5 Sensitivity-Uncertainty Indexes--Simulation of 4 August 1975: Sensitivity to Air-Shed-Wide Peak Ozone Level.. VI-17

VI-6 Hypothetical Case of Needs for Data Acquisition............. VI-19

TABLES

1 Sensivity Measures for Ozone Concentrations above 12 pphm.. ix

2 Sensitivity-Uncertainty Indexes............................ xiv

II-1 Summary of Peak Hourly Averaged Concentrations Observed During the Los Angeles Episode Days Examined in this Study.. II-7

II-2 Summary of Aerometric and Emission Data Bases............. II-8

III-1 Description of the Airshed Model Sensitivity Simulations... III-8

III-2 Air Quality Monitoring Stations in the South Coast Air Basin and the Number of Times They Were Selected for Inclusion in the Reduced Aerometric Data Set........... III-8

III-3 Hydrocarbon Splits by Source Category..................... III-13

III-4 Categories of Emissions Sources and Temporal Distributions... III-25

IV-1 Summary of Maximum Computed and Observed O_3 Concentrations... IV-5

IV-2 Summary of Model Performance in Computing Peak Ozone Concentrations.................................... IV-6

IV-3 Summary of Airshed Model Performance Measures for Ozone for the Los Angeles Simulation.................. IV-8

V-1 Summary of Sensitivity Studies............................. V-10

V-2 Sensitivity Measures for Ozone Concentrations Above 12 pphm (Station Statistics).............................. V-11

V-3 Sensitivity Measures for Ozone Concentrations Above 20 pphm (Station Statistics) V-12

V-4 Sensitivity Measures for Ozone Concentrations Above 12 pphm (Grid Statistics) V-13

V-5 Sensitivity Measures for Ozone Concentrations Above 20 pphm (Grid Statistics) V-14

V-6 Overall Maximum Ozone Levels V-18

V-7 Maximum Ozone Statistics V-19

V-8 Dosages V-20

V-9 Summary of Airshed Model Sensitivity Results V-22

VI-1 Cost Estimates for Data Acquisition VI-7

VI-2 Uncertainty Estimates for Model Input Variables VI-10

VI-3 Sensitivity-Uncertainty Indexes: Sensitivity to Ozone Levels above 12 pphm VI-11

VI-4 Sensitivity-Uncertainty Indexes: Sensitivity to Air-shed-Wide Peak Ozone Level VI-12

VI-5 Hypothetical Cost Estimates for Data Acquisition VI-20

VI-6 Possible Limitations of the Sensitivity Results VI-25

VII-1 Results of the Sensitivity/Uncertainty Analysis for the Los Angeles Basin VII-4

I INTRODUCTION

The Office of Air Quality Planning and Standards (OAQPS) of the U.S. Environmental Protection Agency (EPA) has for some time been examining the feasibility of using grid-based complex atmospheric simulation models that estimate concentration patterns of reactive pollutants for regulatory applications. Atmospheric modeling requires relating changes in emissions to changes in air quality and present grid-based simulation models provide a fairly detailed representation of the spatial and temporal features of atmospheric background species and pollutants. The development and application of such complex models has been considered because of the complexity of the physical and chemical processes governing the formation, transport, and dispersion of secondary* pollutants. Caution in recommending such models, in general, derives from several considerations:

> The limited extent of testing of the predictive performance of these models.

> The imprecision and possible bias associated with their predictions; however, this imprecision and bias has not been shown to be greater--and may well be less--than that associated with the predictions of models presently recommended by EPA for general use. Such recommendations have been conferred only on models used to predict the concentration distributions of inert pollutants.

> The difficulty and cost of acquiring field data (meteorological data, emission inventories, air quality data) used as input information for the model.

> The costs associated with the setup and execution of the models.

---

* Primary pollutants are directly emitted into the atmosphere, whereas secondary pollutants are not directly emitted, but are formed in the atmosphere by the chemical reaction of primary pollutants.

Although evidence related to model performance is limited, it does suggest that uncertainties in model predictions fall well within the bounds of current, accepted practice. For instance, the Systems Applications, Incorporated (SAI) Airshed Model shows relatively little systematic bias for ozone, and computed ozone levels are, on the average, within 25 to 40 percent of the measured values when the measurements exceed the National Ambient Air Quality Standard (NAAQS) of 0.12 ppm. Computed CO and NO_2 concentrations agree with the observations within 40 to 65 percent (Reynolds and Roth, 1980). Of course, reduction in the uncertainties associated with model performance is clearly judged desirable.

The remaining concerns attendant with use of these complex models are cost-related; of the various cost elements, by far the dominant one is that associated with field observations (and data collection) and the compilation of an emission inventory. The difficulty and cost of acquiring such data is perceived to be suficiently great to discourage widespread adoption of grid-based modeling approaches for routine regulatory purposes. Therefore, it appeared necessary to systematically investigate the sensitivity of complex photochemical models to the detail of input information. There are several points to be addressed in such a study.

First, a general methodology must be developed to combine in a rational manner the various components of a definition of model sensitivity to input data. These components are:

> The sensitivity of the model predictions to the level of input data detail considered.

> The uncertainty inherent in the model variables that are related to the input data.*

> The cost of collecting the input data.

Once such a methodology has been developed, it can be used to determine with detail and accuracy which input data must be obtained so that the appropriate emphasis can be established during the development of the data base.

This general methodology must also be applied to a specific example to provide an illustration of its use, as well as a demonstration of its utility. Information on the relative importance of input data obtained from a specific case study can be used to indicate those simplifications that can be made in a complex model without substantially affecting predictions. This type of information is also useful in assessing which input parameters are the most important and which should be considered explicitly in "simple" models. This report describes work carried out to address these issues.

---

* For example, uncertainty in wind direction must be considered in assessing the importance of the amount of meteorological data in model predictions.

An important aspect of this study is the selection of the air quality simulation to be used as a demonstration of the general methodology. Although it would have been desirable to carry out the most general type of sensitivity assessment, it was necessary to restrict the inquiry in several important ways:

> The selection of one particular grid-based, complex air quality simulation model--the SAI Airshed Model. This model was chosen by OAQPS because it is presently being evaluated in several U.S. cities. Clearly, however, another grid-based simulation model could have been used.

> The selection of the Los Angeles area as a prototype. This choice was based primarily on considerations of cost and timeliness. Evaluation of model performance had been completed for a suitable test day for Los Angeles; this work would not have to be repeated, and sensitivity studies could commence without undue delay. It was acknowledged, however, that the study of one to three additional cities, located in other geographical areas and having different climatologies, would be highly desirable.

> The selection of particular days for study. Because each day has its unique meteorological and air quality characteristics, the days selected may not typify other days of interest--i.e., "bad" or "worst" days under various meteorological regimes. However, two reasonably representative bad (i.e., high ozone) days were chosen for study--in fact, these days were selected as prototypical in at least two prior research studies as well.

> The study of a fifteen-hour simulation period. Unbiased sensitivity estimates can be derived only from multiple-day simulations, as usually emissions (and emission changes) from prior day(s), as well as emissions from the present day, influence the simulation results. Thus, one-day simulations underestimate sensitivity to changes in emissions and overestimate sensitivity to initial conditions. Multiple-day simulations were not carried out during the course of this study because of

- The unavailability of a tested base case.

- Lack of adequate experience with such simulations.

- Cost restrictions.

> The limited range of applicability of sensitivity estimates. Calculated "sensitivities" apply only for the perturbations studied; the effects of other perturbations or variations of differing magnitudes can not be directly extrapolated because of the nonlinear nature of the model.

Despite these limitations, it is worthwhile to undertake such a sensitivity analysis for the following reasons:

> The present lack of any useful sensitivity results or analyses pertaining to data base degradation.

> The potential usefulness of sensitivity results in defining relative priorities in data needs.

> The need for information of this type in estimating the overall costs of photochemical grid model use.

> The inferences that might be drawn about the utility of simple models that do not explicitly use several of the input parameters required by the airshed model.

> The commonality of meteorological conditions in varied geographical areas that lead to the highest observed concentrations of ozone.

> The ability to bring available information and knowledge to bear in modifying the findings, through qualitative arguments, so that they apply more generally.

Hence, the sensitivity analysis was planned and carried out.

In chapter II of this report, certain important characteristics of the Los Angeles Basin as they relate to the formation of photochemical oxidant are discussed and the data resources that were available to support the present study are summarized. Chapter III introduces the sensitivity simulations and provides the rationale behind the selection of each and the methodology employed in preparing model input files. Next, the two base case simulations are discussed. Evaluation of model performance in simulating ozone concentrations on the days of 26 June 1974 and 4 August 1975 is presented in chapter IV.

The results of 22 sensitivity runs are discussed in chapter V. Certain measures that are helpful in quantifying model sensitivity are introduced and then applied to these results. Overall sensitivities are summarized in tabular and graphical form. Detailed results of the individual simulations are compiled in appendix A of this report.

The general methodology used to evaluate the relative importance of input data for a complex air quality model is developed and applied in chapter VI. The chapter begins with the definition of a sensitivity-uncertainty index. After estimating costs for acquisition of various types of data, the index and cost estimates are applied, yielding a ranking of data needs. The chapter concludes with a discussion of the specificity, generalizability, and limitations of the results.

Chapter VII summarizes the principal findings and conclusions. Appendix B discusses model input requirements, the level of detail in data bases currently available throughout the United States, and results of recent photochemical grid model sensitivity studies.

II CHARACTERISTICS OF THE SOUTH COAST AIR BASIN

This chapter discusses the selection of the South Coast Air Basin (SCAB) as the urban area prototype for the airshed model sensitivity study. Unique features of the SCAB, as they relate to oxidant formation processes and the preparation of a photochemical modeling data base, are considered.

A. STUDY AREA SELECTION

At the inception of this study, two urban areas possessed photochemical modeling data bases of sufficient "richness" to justify their selection for subsequent sensitivity analyses. These data bases were for the Los Angeles, California and St. Louis, Missouri, urban complexes.

Determination of the most suitable study area involved a number of trade-offs and consideration of previous experiences in oxidant episode simulation. For Los Angeles, for example, detailed mobile, area, and stationary point source emission inventories were readily available in a format well-suited to modification. (Modification to the baseline inventory would later be made to reflect reduced levels of detail in the information data base used to compile an emission inventory.) The St. Louis inventory, in contrast, did not offer an equivalent degree of flexibility in altering the emission rates in various source categories. Indeed, it was unclear whether any changes to the St. Louis mobile source inventory could be made (e.g., reflecting outdated VMT estimates) without a costly and time consuming effort to obtain the relevant emission files.

The physical setting and meteorology of the South Coast Air Basin presented another trade-off situation. The meteorology of this area is more complex than that of many other urban areas in the United States for which photochemical modeling may be contemplated. One consequence in selecting Los Angeles, then, is the possibility that the need for detailed meteorological data may become overstated (relative to other data needs) as a result of its complex meteorology.

Previous experience in simulating oxidant episodes in Los Angeles was also a factor in determining which of the two urban areas should be the focal point for the study. Adequate airshed model performance in simulations of the South Coast Air Basin had been demonstrated in previous studies (e.g., Tesche and Burton, 1978; Reynolds et al. 1979). At the time the decision was made as to which urban area was to serve as the prototype, model performance evaluation studies in St. Louis were still underway. Establishment of a suitable base case appeared to be a less costly endeavor for Los Angeles than for St. Louis. Furthermore, a number of previous airshed model sensitivity runs had been carried out already with the Los Angeles data base, and these studies could be used to supplement the findings of the present work. Based upon the above considerations, the South Coast Air Basin was chosen for the sensitivity analysis.

The following discussion provides an overview of the more important oxidant formation processes in the Los Angeles basin, for the reader who may be unfamiliar with certain unique features of the area.*

B. PHOTOCHEMICAL OXIDANT FORMATION IN THE SOUTH COAST AIR BASIN

Important elements in the formation of oxidant in the South Coast Air basin are the unique meteorology, complex topography, and varied spatial and temporal distribution of precursor emissions. In the presence of an intense summer sun, these variables assume values that produce frequent high concentrations of photochemical oxidants. Two major meteorological mechanisms--a persistent low-level subsidence inversion and a thermally driven sea breeze circulation--contribute to the accumulation and transport of both primary pollutants and secondary photochemical oxidants. Topography influences the transport of polluted air masses. Because of wind channeling by the surrounding mountainous terrain, certain receptor areas (e.g., Pasadena, Upland, Fontana) often experience high oxidant levels. Also, a varied spatial distribution of precursor emissions, particularly influenced by temporally varying automotive exhaust, combines with the two physical mechanisms mentioned above to form a dense chemically reactive pollutant cloud, thereby creating a potential for high of oxidant concentrations.

The following subsections elaborate on the roles of meteorology, emissions, and topography in the oxidant formation process. The interaction of these physical and chemical mechanisms leads to the typical daily oxidant formation pattern in the SCAB during the summer-to-autumn smog season.

---

* Figure IV-5 on page IV-19 identifies important landmarks and cities in the South Coast Air Basin.

1. Meteorology: The Inversion and the Sea Breeze

Formation of oxidant in the basin during the months of May through October (the smog season) is dominated by the intensity and height of a persistent low-level temperature inversion and by the strength of the land-sea breeze. Southern California weather is strongly influenced by the semi-permanent east Pacific high pressure system. Throughout the summer, this intense high pressure system generates anticyclonic flow aloft diverging toward the southeast. This divergence aloft creates air mass subsidence, thus creating an intense inversion as the sinking air heats through adiabatic compression. The warm sinking air is extremely bouyant when contrasted to the layer of cool moist marine air beneath the inversion. A stable situation results; the inversion acts as a virtual "lid" that both defines and restricts the size of the cooler marine layer.

The summer inversion over the SCAB suppresses the vertical dispersion of pollutants, causing emissions to become "trapped" in a shallow layer the size of which is defined by the inversion height. The size of the marine layer varies with time and location; that is, its depth at a particular location generally increases during the day. In addition, as the layer moves inland its depth increases. Although inversion base heights range from an average of 450 meters in the morning to 750 meters in the afternoon (AQMD, 1979), the most severe cases of oxidant formation occur when the base height [measured by the 0700 PST Los Angeles International Airport (LAX) rawinsonde sounding at the coast] lies between 250 and 500 meters and the inversion strength (measured as the top temperature minus the base temperature) is greater than 10° C (Zeldin et al., 1976).

A major component of pollutant dispersion in the SCAB is the thermally forced sea breeze. The sea breeze is generated from the temperature contrast between warm inland areas and the relatively cool coastal waters of the Pacific Ocean. The sea breeze commences midmorning and moves inland during the day, persisting into the early evening. As evening temperatures diminish inland, the thermal gradient reverses, and a weak offshore flow is established. Several trajectory studies of this sea breeze system (Neiburger, 1969; Holmes et al., 1956; Drivas and Shair, 1974; and Angell, Dickson, and Hoecker, 1976) have verified this persistent pattern. In general, the direction and strength of the sea breeze determines the potential for dilution of polluted air masses. Under typical summertime conditions, morning emissions of precursors (say from the 0600 to 0900 PST period) accumulate in the metropolitan and coastal source areas under generally stagnant winds. The ensuing sea breeze gathers strength and advects the pollutant cloud inland.

The daytime pollutant transport pattern usually results in peak oxidant concentrations in the early afternoon at downwind receptor sites in the eastern half of the basin with secondary maxima in the San Fernando Valley. As midafternoon temperatures increase, surface heating weakens or destroys the inversion, causing trapped oxidants to mix within a deeper layer, thereby diluting surface concentrations. In contrast, on days when the sea breeze is weak, transport and dilution of the pollutant cloud are negligible. Under severe conditions, a weak offshore flow occurs during the day, confining the polluted cloud over the downtown metropolitan area. These conditions often coincide with an intense low-level subsidence inversion, resulting in several days of high concentrations of photochemical smog (Zeldin and Cassmassi, 1979).

2. Emission Distribution and Oxidant Formation

The spatial and temporal distributions of precursor emissions bear directly on the formation of oxidant in the basin. The primary source of precursor emissions in the SCAB is automobile traffic. Currently, over 9 million people live in the basin, approximately 7 million of these in Los Angeles County (AQMD, 1979). Because of limited mass transit facilities, the automobile dominates commuter transit, resulting in high levels of HC and NO_x emissions during the morning and evening rush hours. 1974 traffic contributions to precursor emissions are estimated at approximately 44 percent of the hydrocarbon inventory and 51 percent of the total NO_x emissions. Estimated 1987 emissions from mobile sources are 24 percent and 37 percent, respectively (Souten et al. 1980).

Although the spatial distribution of the traffic emission inventory varies with the time of the day, traffic emissions are concentrated generally around downtown Los Angeles and along the major traffic arteries leading into the metropolitan area. Emissions from traffic increase during early morning rush hours (0600 to 0900), a period typically characterized by light stagnating winds and a low-level inversion. Simultaneously, contributions from several point sources, particularly power plants (NO_x) and refineries (NO_x and HC) clustered along the coast from Lennox to Costa Mesa, add to the spatially varied distribution of precursor emissions.

Although a significant number of these sources are elevated, most precursor emissions fail to penetrate the inversion; rather, they remain within the marine layer. From the combination of the traffic and point sources, precursor emissions begin to accumulate, forming a large reactive pollutant cloud. With time, the reactive cloud "ages," and chemical transformation of primary pollutants takes place. Concurrently, the pollutant cloud is advected inland by the sea breeze. Although oxidant production is evident in the polluted cloud, high levels are not usually observed in the metropolitan area (downtown Los Angeles). This is due in part to NO_x scavenging by continuous automotive emissions. In addition, by midday the sea breeze has advected relatively clean marine air over the west-central basin.

3. Effects of Topography on Pollutant Transport

Transport of photochemical oxidants in Los Angeles is largely constrained by the topography of the basin. Elevated terrain, including the San Gabriel, Santa Monica, and San Bernardino mountains, and the Puente Hills, influence local wind patterns. Although the mountains virtually surround the basin, they do not present an impenetrable barrier blocking dispersion of regional pollutants.* A major effect of the topographic relief is the channeling of wind flow, which forces pollutants to converge in some areas (e.g., the San Fernando Valley) and to disperse in other areas (southeast of Chino). As a consequence, three frequent transport patterns are observed commonly.

The dominant oxidant transport pattern in the SCAB follows a course from coastal source areas inland to the San Gabriel Valley. The route is directly related to the east-west ridge line transversing the San Gabriel and Santa Monica mountains. In effect, the orientation of the ridge line and the basic direction of the sea breeze coincide, thus enhancing transport of the polluted cloud inland. The channeling of high concentrations of oxidant and precursors in the direction of Pasadena, Fontana, and San Bernardino is verified by the occurrence of persistently high concentrations of oxidant in the east basin (Davidson et al., 1979).

In contrast, Riverside is an east basin inland receptor that is affected primarily by pollutant transport through the Santa Ana Canyon. When offshore flow forces the major pollutant cloud into a more southerly transport trajectory, Riverside can experience some of the highest oxidant concentrations in the SCAB (Zeldin and Cassmassi, 1979).

The San Fernando Valley represents an alternate receptor for the polluted cloud. Areas of the east San Fernando Valley (e.g., Burbank) frequently experience higher values of oxidant and precursor concentrations, than receptors in the western portion of the valley. High levels of oxidants can be observed when southerly flow exists over Southern California (Zeldin and Cassmassi, 1979).

---

* Although a general accumulation of photochemical oxidants and precursors is observed along the foothills of the San Gabriel Mountains, intense solar heating of the mountaintops will usually generate an upslope wind system of sufficient strength to penetrate the inversion and to vent surface pollutants aloft.

4. The SCAB as a Case Study

Oxidant formation in the SCAB represents a unique situation in which routinely monitored high concentrations are directly related to the interaction of the three principal component mechanisms: meteorology, topography, and emissions. Although other metropolitan areas (e.g., the Northeast megalopolis, Albuquerque, and Denver) experience periods of high oxidant concentrations, the severity of oxidant formation and the conditions affecting it are different from those observed for the SCAB. For example, the latitude in Albuquerque is similar to that of Los Angeles and, as a result, these cities experience approximately the same solar input. However, Albuquerque is not subject to the effects of a low-level subsidence inversion or to a similar intensity of precursor emissions. Consequently, oxidant episodes are less frequent and severe, compared with those in Los Angeles. Although the emission density and frequency of stagnant summer periods for the Northeast megalopolis (including Washington, D.C., Philadelphia, and the tristate region of New York, New Jersey, and Connecticut) parallel those of the SCAB, stagnation in the Northeast is not exacerbated by the presence of an intense subsidence inversion.

The SCAB also illustrates a direct cause-effect situation. Unlike the case for the Northeast, where precursor emissions in Philadelphia may affect oxidant concentrations observed in New York, oxidant formation in Los Angeles is largely attributed to local emissions; "upwind" pollutant concentrations typically reflect background concentration levels attributable to previous emissions from the area.

In summary, the SCAB represents an interesting case study area for conducting photochemical modeling and sensitivity analyses and for examining specific features of the oxidant formation process. However, as discussed later, certain unique features of the basin make it necessary to exercise care in generalizing the sensitivity study results to other urban areas, where the importance of various physico-chemical oxidant forming processes may differ.

C. DESCRIPTION OF TWO LOS ANGELES OXIDANT EPISODES

The airshed model sensitivity simulations are based on two ozone episodes[*] that occurred in Los Angeles during the summers of 1974 and 1975. During the 26 to 29 June 1974 episode, a land-sea breeze regime prevailed--a common occurrence in the SCAB (Kieth and Selik, 1977). At night, pollutants were advected out over the Santa Monica Bay; with the ensuing morning sea breeze, some of this material was reintroduced into the air basin. Such conditions are representative of a "typical" smoggy day in Los Angeles. On 4 August 1975, during another high ozone episode, onshore winds prevailed throughout the entire 24-hour period. Analysis of the synoptic meteorology and local wind station measurements revealed persistent onshore winds for the two days preceding 4 August, as well. Selection of this day allows us to examine a set of meteorological conditions that are characteristically different from those occurring between 26 and 29 June 1974. This, in turn, provides some insignt into how closely tied results of the sensitivity analyses are to a given set of meteorological conditions. Table II-1 summarizes the peak pollutant concentrations recorded during the two episodes.

TABLE II-1. SUMMARY OF PEAK HOURLY AVERAGED CONCENTRATIONS OBSERVED DURING THE LOS ANGELES EPISODE DAYS EXAMINED IN THIS STUDY

	1974		1975
Pollutant	26 June	27 June	4 August
NO_2	26 pphm	36 pphm	19 pphm
O_3	34 pphm	49 pphm	32 pphm
CO	13 ppm	14 ppm	8 ppm
Sulfate	7.2 $\mu g/m^3$	No data	28.9 $\mu g/m^3$
SO_2	10 pphm	17 pphm	12 pphm

D. AVAILABLE DATA RESOURCES

Two of the major reasons the Los Angeles basin was chosen for this study are the existence of an extensive meteorological and air quality monitoring network and the recent compilation of a detailed emission inventory. Table II-2 summarizes the extent of the data available for

---

[*] The rationale for selecting two episodes instead of one is discussed in the next chapter.

TABLE II-2. SUMMARY OF AEROMETRIC AND EMISSION DATA BASES

(a) The 1974 Los Angeles Data Base

Type of Data	Nature of Data Collected	Remarks	Source or * Organization(s) Collecting Data	Reference
Air quality data				
Surface pollutant concentrations	39 air quality monitoring stations	Data are hourly averaged values except sulfate, which is reported as a 24-hour average	Various APCDs	Ross (1975)
Pollutant concentrations aloft	None			
Meteorological data				
Surface winds	60 surface wind monitoring stations	Data are hourly averaged and instantaneous values	APCDs, ARB, NWS, PMTC	Ross (1975)
Upper Level winds	Radiosondes at Riverside (0600 PST), El Monte (0700, 1330), LAX (0630, 1230), PT. Mugu (0400, 1000, 1600 PST) and San Nicholas Island (1400 PST)		NWS, ARB, PMTC	Ross (1975)
Mixing Depths	Acoustic sounder at El Monte	Data are hourly averaged values	ARB	Ross (1975)
	Aircraft spiral at Riverside (0600 PST)		ARB	Ross (1975)
	Radiosondes at LAX (0630, 1230 PST), El Monte (0700, 1300 PST), PT. Mugu (0400, 1000, 1600 PST), and San Nicholas Island (1400 PST)			

TABLE II-2 (Continued)

Type of Data	Nature of Data Collected	Remarks	Source or * Organization(s) Collecting Data	Reference
Surface temperature	Temperatures at several stations	Data are hourly averaged values		
Solar radiation	Radiation measurements at LAX, downtown Los Angeles, and Riverside	Data are hourly averaged values	ARB, APCDs	
Humidity	Humidity at several stations	Data are hourly averaged values	NWS, APCDs	
Visibility	Visibility at 7 airports	Data are instantaneous values recorded hourly	NWS	
Cloud cover	Cloud cover at 7 airports	Data are instantaneous values recorded hourly	NWS	
Emission data				
Traffic	Emission estimates derived from the LARTS transportation model and the ARB DTIM emissions	Peak, offpeak, and total emission rates for THC, SO_2, NO_x, and CO are available. Percentages of hot and cold starts and hot soaks are included in the inventory with spatial and temporal resolution. Vehicle speed distributions and vehicle types (4 classes) are included	ARB	
Refinery	Emission estimates for NO_x and reactive hydrocarbons	Emissions are based on nominal daily rates		SCAQMD (1978)
Aircraft	Gridded emission estimates for 34 airports			SCAQMD (1978)
Power plants	18 power plants for NO_x, SO_2, organics, particulates	Diurnal distribution of emissions based on inspection of daily operating logs	Utilities	SCAQMD (1978)
Distributed area sources	Emission estimates for RHC and NO_x			SCAQMD (1978)
Other stationary sources	Emission estimates for RHC and NO_x			SCAQMD (1978)

TABLE II-2 (Continued)

(b) The 1975 Los Angeles Data Base

Type of Data	Nature of Data Collected	Remarks	Source or Organization(s) Collecting Data*	Reference
Air quality data				
Surface pollutant concentrations	48 air quality monitoring stations	Data are hourly averaged except sulfate, which is reported as a 24-hour-average	SCAQMD	Tesche and Burton (1978)
Pollutant concentrations aloft	None			
Meteorological data				
Surface winds	56 surface wind monitoring stations	Data include hourly averaged and instantaneous values	APCDs, ARB, NWS, PMTC	Tesche and Burton (1978)
Upper level winds	Radiosondes at LAX (1130 PST) El Monte (0610, 1230 PST), Pt. Mugu (0350, 1004, and 1615 PST), and San Nicholas Island (1405 PST)			
	Pibals at Burbank (1940 PST) and San Bernardino (0345 and 1030 PST)			
Mixing depths	Aircraft spiral at Riverside (0300 PST)		ARB, NWS, PMTC	Tesche and Burton (1978)
	Radiosondes at LAX (0530, 1030 PST), El Monte (0610, 1230 PST, Pt. Muqu (0350, 1004, 1515 PST), and San Nicholas Island (1405 PST)			
Surface temperatures	Temperatures at 15 stations	Data are hourly averaged values	NWS, SCAQMD	
Solar Radiation	Radiation measurements at El Monte and Riverside	Data are hourly averaged and daily averaged values	NWS, SCAQMD	
Humidity	Humidity at 15 stations	Data are hourly averaged values	NWS, SCAQMD	

TABLE II-2 (Concluded)

Type of Data	Nature of Data Collected	Remarks	Source or Organization(s) Collecting Data*	Reference
Visibility	Visibility at 7 airports and 2 other locations	Data are instantaneous values recorded every hour	NWS, SCAQMD	
Cloud Cover	Cloud cover at 7 airports and 2 other locations	Data are instantaneous values recorded every hour	NWS, SCAQMD	
Emission data				
Traffic	Emission estimates derived from the LARTS transportation model and the ARB FWY011 emission model	Peak, offpeak, and total emission rates for THC, SO_2, NO_x, and CO are available. Percentages of hot and cold starts and hot soaks are included in the inventory with spatial and temporal resolution. Vehicle speed distributions and vehicle types (4 classes) are included	ARB	
Refinery	Emissions for organic gases, NO_x, SO_2, CO, and particulates estimated for 28 facilities		SCAQMD	SCAQMD (1978)
Aircraft	Gridded emission estimates for 34 airports		SCAQMD	SCAQMD (1978)
Power plants	Emission estimate of NO, NO_2, SO_2, $SO_4^=$, and CO for 13 plants	Diurnal distributions of emissions based on inspection of daily operating logs	Utilities	SCAQMD (1978)
Distributed area sources	Emission estimates for RHC and NO_x		Utilities	SCAQMD (1978)
Other stationary sources	Emission estimates for organic gases, NO_x, SO_2, CO, and particulates		Utilities	SCAQMD (1978)

---

* PMTC = Pacific Missile Test Center, APCD = Air Pollution Control District, ARB = Air Resource Board
NWS = National Weather Service, SCAQMD = Southern California Air Quality Monitoring District.

photochemical modeling of the June 1974 and August 1975 episodes. Important aspects of the data base are mentioned briefly.

1. Emission Inventory

The 1975-1976 emission inventory for the SCAB is a result of the Air Quality Management Plan (AQMP) development process initiated in 1976 by the Southern California Association of Governments (SCAG) and the South Coast Air Quality Management District (SCAQMD). The development of the emission inventory was separated organizationally and procedurally into several components: stationary sources, area sources and off-road vehicles, and mobile sources (Souten, Anderson, and Ireson, 1980). The stationary source inventory development was initiated at the SCAQMD using their files on source activity. Much of the work performed at the district involved manual transcription of emission records and inventory reformatting necessitated by the model. District data were then modified by a contractor to approximate average summer day emissions.

Area source and off-road motor vehicle inventories were developed at SCAG, primarily using demographic data, whereas the on-road mobile source inventory was generated by starting with SCAG demographic and transportation system data. These data were used as inputs to the California Department of Transportation (Caltrans) traffic modeling system, which in turn provided the necessary inputs for the Direct Travel Impact Model (DTIM), a model provided by the California Air Resources Board (CARB) to generate actual mobile emissions according to vehicle fleet mix, hot and cold starts, hot soaks, and so on. The ARB was also involved in assembly of the inventory components through a system designed to allow automated generation of gridded emissions rates as required by photochemical simulation models. Details of the 1975-1976 emission inventory are discussed in the SCAB Non-Attainment Plan (NAP) and its supporting working papers, prepared jointly by SCAG and SCAQMD.

2. Aerometric Data

As table II-2 indicates, meteorological and air quality data for the SCAB are available from several sources. Data from these sources were acquired in the course of previous applications of the airshed model to the basin (Reynolds et al., 1979; Tesche and Burton, 1978).

One element of the aerometric data base not listed in table II-2 warrants special mention. Routine reporting of hydrocarbon concentration data in the basin includes total hydrocarbons (THC) and non-methane hydrocarbons (NMHC), the latter obtained primarily from the difference between THC and methane determinations. The data are reported to the nearest part per million. Lack of precision in the reactive hydrocarbon data (due to rounding to the nearest ppm) was partially ameliorated by a special reactive hydrocarbon monitoring program sponsored by the CARB during the summer months of 1974 and 1975. From June to September in both 1974 and 1975, a program was conducted to determine the nature and extent of non-methane hydrocarbons through C_{10} in the atmosphere of the SCAB. Sampling sites selected for the study were Long Beach, downtown Los Angeles, El Monte, Azusa, and Upland. Atmospheric samples were collected using automatic sequential samplers at the following times:

- 0200-0500 hours (pre-dawn).
- 0600-0900 hours (peak traffic).
- During a three hour interval corresponding to peak oxidant occurrence at each site.

Total hydrocarbon monitoring was also performed, thereby enabling statistical correlations to be drawn between ambient THC and RHC concentration levels.

E. Previous Sensitivity Studies in the South Coast Air Basin

Several studies of photochemical grid model sensitivity to variation in input parameters have been reported in the literature in the last several years. Most of these studies have involved application of the SAI Airshed Model or the LIRAQ model developed by Lawrence Livermore National Laboratory to various urban areas (i.e., Los Angeles, San Francisco, Denver, St. Louis, Sacramento). Reported sensitivity analyses focusing on the South Coast Air Basin have been performed almost exclusively with the Airshed Model. Appendix B of this report discusses the principal findings of sensitivity studies conducted for the South Coast Air Basin, as well as for other urban areas.

III STRUCTURE OF THE SENSITIVITY SIMULATIONS

This chapter introduces each of the sensitivity simulations and provides the rationale behind the reduction in the level of detail of information available for preparing airshed model inputs. Appendix A compares in detail the changes in input files for sensitivity runs and base case runs. Chapter V discusses what impacts these changes have on predicted oxidant levels.

A. GENERIC CLASSIFICATION OF SENSITIVITY RUNS

The 22 sensitivity runs performed in this study may be grouped into the following general categories:

- Meteorology
- Air Quality
- Emissions
- Model Structure, (i.e., grid resolution, number of vertical layers).

Table III-1 describes each simulation briefly. In a following section the input preparation procedures for each run are summarized. First, however, a few procedural issues that bear on the overall structure of the sensitivity study need to be identified.

B. PROCEDURAL ISSUES

1. Modeling Region

Two major considerations governed the selection of the modeling region. First, substantial computing cost reductions could be gained by limiting the size of the region (by eliminating grid cells over the ocean, mountains, or desert, for example). The second consideration was that the

TABLE III-1. DESCRIPTION OF THE AIRSHED MODEL SENSITIVITY SIMULATIONS

Run Number	Description	Characteristics of the Base Case		Characteristics of the Sensitivity Run
1	74 upper air	Upper level wind	Radiosondes at Riverside (0600 PST), El Monte (0700, 1330), LAX (0630,1230), Pt. Mugu (0400, 1000, 1600 PST) and San Nicholas Island (1400 PST)	26 June 1974 wind fields and mixing depths based on 0700 and 1330 El Monte radiosondes
		Mixing depths	Acoustic sounder at El Monte	
			Aircraft spiral at Riverside (0600 PST)	
			Radiosondes at LAX (0630, 1230 PST), El Monte (0700, 1300 PST), Pt. Mugu (0400, 1000, 1600 PST), and San Nicholas Island (1400 PST)	
2	75 upper air meteorological data	Upper level winds	Radiosondes at LAX (1130 PST), El Monte (0610,1230 PST), Pt. Mugu (0350, 1004, and 1615 PST), and San Nicholas Island (1405 PST)	4 August 1975 wind fields and mixing depths based on 0610 and 1230 El Monte radiosondes
			Pibals at Burbank (1940 PST) and San Bernardino (0345 and 1030 PST)	
		Mixing depths	Aircraft spiral at Riverside (0300 PST)	
			Radiosondes at LAX (0630, 1030 PST), El Monte (0610, 1230 PST), Pt. Mugu (0360, 1004, 1615 PST), and San Nicholas Island, (1405 PST)	
3	74 surface and upper air meteorological data	Surface winds	60 surface wind monitoring stations	26 June 1974 wind fields and mixing depths based on 0700 and 1330 El Monte radiosondes and surface winds and temperatures at ten monitoring stations*
		Upper level winds	Radiosondes at Riverside (0600 PST), El Monte (0700, 1330), LAX (0630, 1230), Pt. Mugu (0400, 1000, 1600 PST) and San Nicholas Island (1400 PST)	
		Mixing depths	Acoustic sounder at El Monte aircraft spiral at Riverside	
			Radiosondes at LAX (0630, 1230 PST), El Monte (0700, 1300 PST), Pt. Mugu (0400, 1000, 1600 PST), and San Nicholas Island (1400 PST)	
		Surface temperatures	Temperatures at several stations	

TABLE III-1 (Continued)

Run Number	Description	Characteristics of the Base Case		Characteristics of the Sensitivity Run
4	75 surface and upper air meteorological data	Surface winds	56 surface wind monitoring stations	4 August 1975 wind fields and mixing depths based on 0610 and 1230 El Monte radiosondes and surface winds and temperatures at ten monitoring stations
		Upper level winds	Radiosondes at LAX (1130 PST), El Monte (0610, 1230 PST), Pt. Mugu (0350, 1004, and 1615 PST), and San Nicholas Island (1405 PST)	
			Pibals at Burbank (1940 PST) and San Bernardino (0345 and 1030 PST)	
		Mixing depths	Aircraft spiral at Riverside (0300 PST)	
			Radiosondes at LAX (0630, 1030 PST), El Monte (0610, 1230 PST), Pt. Mugu (0350, 1004, 1615 PST) and San Nicholas Island (1405 PST)	
		Surface temperatures	Temperatures at 15 stations	
5.1	Ground level initial and boundary conditions	Initial and boundary conditions based on synthesis of data from 39 air quality monitoring stations		26 June 1974 initial and boundary conditions based on 10 surface monitoring stations
5.2	74 reduced meteorological and air quality data	See descriptions of runs 3 and 5.1 above		26 June 1974 wind fields, mixing depths, initial conditions, and boundary conditions based on 0700 and 1300 El Monte radiosondes and surface winds, temperatures, and pollutant concentration measurements at ten monitoring stations
6	Upper air quality data	Upper air concentration profiles assumed to decrease with height to assumed background values		26 June 1974 simulation with layer of precursor material aloft, estimated from 3-D gradient study profiles
7	75 clean air boundary conditions	Boundary conditions based on synthesis of data from 48 air quality monitoring stations		4 August 1975 simulation with less precursor material aloft; clean air boundary conditions
8	74 clean air boundary conditions	Boundary conditions based on synthesis of data from 39 air quality monitoring stations		26 June 1974 simulation with less precursor material aloft; clean air boundary conditions
9	74 assumed RHC/NO_x ratio	Reactive hydrocarbon concentrations based on empirical relationship derived from ambient RHC measurements by Mayersohn et al. (1975) in 1974		26 June 1974 simulation with initial hydrocarbon concentrations based on an assumed aerometric hydrocarbon/NO_x ratio of 7; clean air boundary conditions
10	75 assumed RHC/NO_x ratio	Reactive hydrocarbon concentrations based on empirical relationship derived from ambient RHC measurements by Mayersohn et al. (1976) in 1975		4 August 1975 simulation with initial hydrocarbon concentrations based on an assumed aerometric hydrocarbon/NO_x ratio of 7; clean air boundary conditions
11	74 hydrocarbon emissions speciation	Volatile organic compound speciation based on research by KVB Engineering, Incorporated, under EPA Contract No. 68-02-3209		26 June 1974 reactive hydrocarbon emissions split according to an assumed composite urban mix

TABLE III-1 (Concluded)

Run Number	Description	Characteristics of the Base Case	Characteristics of the Sensitivity Run
12	75 hydrocarbon emission speciation	Volatile organic compound speciation based on research by KVB Engineering, Incorporated, under EPA Contract No. 68-02-3209	4 August 1975 reactive hydrocarbon emissions split according to an assumed composite urban mix
13	74 mobile source	Emission estimates derived from the LARTS transportation model and the ARB DTIM emission model	26 June 1974 mobile source emissions based on outdated transportation model and emission simulator
14	74 gas sales	Emission estimates derived from the LARTS transportation model and the ARB DTIM emission model	26 June 1974 mobile source emissions based on estimated vehicle fuel economy, areawide fuel sales, ensemble emission factors, and composite urban vehicle mix
15	75 gas sales	Emission estimates derived from the LARTS transportation model and the ARB DTIM emission model	4 August 1975 mobile source emissions based on estimated vehicle fuel economy, areawide fuel sales, ensemble emission factors, and composite urban vehicle mix
16	Point source	Day-specific, hourly-average emission rates for power plants obtained directly from the utilities	26 June 1974 power plant emissions based on annual average emission rates and average diurnal generating profile
17	74 area source	Detailed, distributed area source inventory developed by the South Coast Air Quality Management District (SCAQMD, 1978)	26 June 1974 distributed area sources spatially allocated according to the demographic distribution
18	75 area source	Detailed, distributed area source inventory developed by the South Coast Air Quality Management District (SCAQMD, 1978)	4 August 1975 distributed area sources spatially allocated according to the demographic distribution
19	Temporal resolution	Detailed, distributed area source inventory developed by the South Coast Air Quality Management District (SCAQMD, 1978)	26 June 1974 area sources temporally resolved according to assumed generic diurnal profiles
20	10 km	Horizontal grid resolution of 5 km	26 June 1974 simulation based on a 10-km horizontal grid mesh
21	two-layer model	Four levels of grid cells	26 June 1974 simulation with one layer below the inversion and one above
22	one-layer model	Four levels of grid cells	26 June 1974 simulation with one layer below the inversion

---

* The ten stations are: downtown Los Angeles, Azusa, Burbank, Long Beach, Anaheim, San Bernardino, Reseda, Lennox, Prado Park, and Upland.

region should not be made too small. If the modeling region was defined as only a small fraction of the overall urban area, thereby excluding certain monitoring stations, emission sources, and so forth, it is possible that the importance of a particular model input could be missed. Figure III-1 defines the modeling region. It encompasses nearly all of the major emission sources in the SCAB and has sufficient coverage of the Santa Monica Bay and Pacific Ocean to contain aged pollutants advected westward with the nocturnal land breeze.

2. Multiple Base Cases

The influence of different meteorologies on the sensitivity study results was of concern from the outset of the study. A Los Angeles application would seem to provide a more extreme test of model sensitivity to the availability of input data. Reduction in the number of wind monitors in St. Louis, for example, would probably not degrade the resultant wind field as much as would a corresponding reduction in Los Angeles wind monitoring stations. Experience in preparing mixing depths for St. Louis generally indicates no substantial difference between the estimated mixing depths at each of the reporting sites. In Los Angeles, the close proximity of coastal and desert environs leads to shallow mixing depths (from 250 to 300 meters) along the coastal margin and to deep mixing depths (from 900 to 1100 meters) farther inland during the midday hours. To investigate whether the selection of different meteorological episodes might materially alter subsequent ranking of data needs, both the 26 June and 4 August episodes were chosen as base cases.

3. Simulation Duration

To keep overall computation costs to an acceptable level, the simulation duration was limited to less than one day. The implications of this constraint for the meaning and generality of the sensitivity results are discussed in chapter VI. Analysis of the ozone concentration levels for 26 June and 4 August indicated that nearly all monitoring stations observed the peak concentrations well before 1800 PST. Therefore, 1900 PST was selected as the time to terminate the runs. The simulation starting time was selected after examination of the emission inventory; at 0400 emissions are relatively low compared to the levels during the ensuing morning rush-hour. It was felt that starting at 0400 would not bear adversely on the sensitivity simulation involving temporal redistribution of emissions.

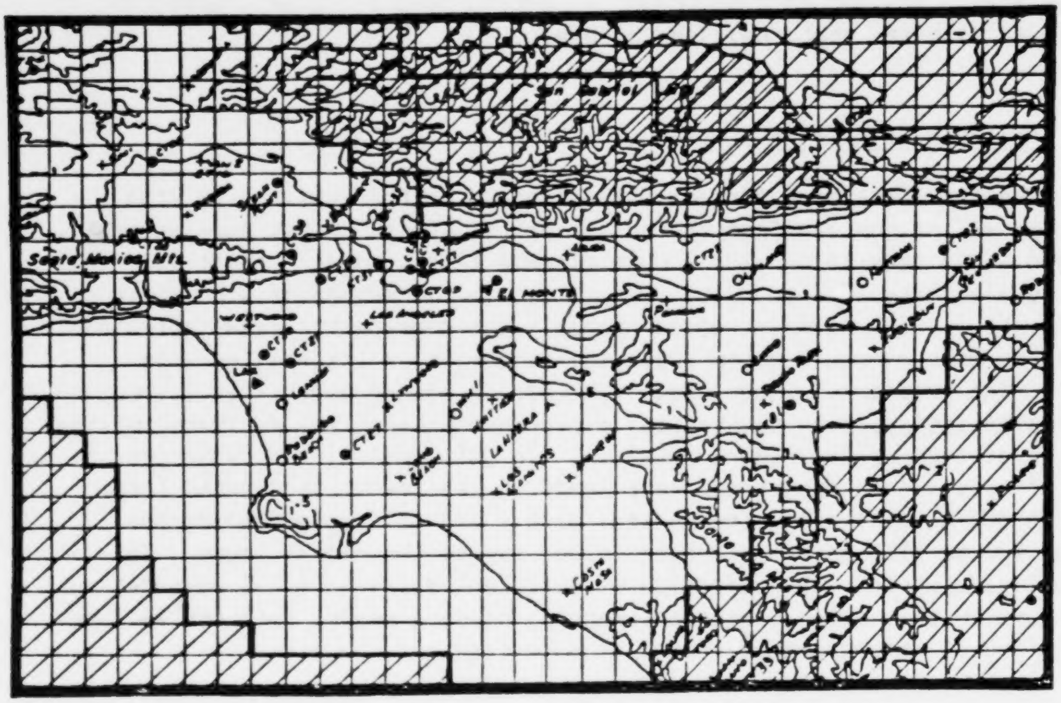

Santa Monica Mtn
EL MONTE
WESTWOOD
LOS ANGELES
LA HABRA
CT93

FIGURE III-1. MODELING REGION USED IN THE AIRSHED MODEL SENSITIVITY STUDY.
(Calculations not performed for the crossed areas.)

4. Selection of the Reduced Set of Air Quality Monitors

As discussed in the following section, certain simulations involve the use of a limited set of surface aerometric monitors. The reduced set of 10 surface air quality monitors was selected by the following methodology. Eight SAI technical staff members with individual backgrounds in meteorology, atmospheric chemistry, emission inventory preparation, aerometric monitoring, modeling, and so on were asked to identify on figure III-1 the locations of 10 sites at which they would locate aerometric monitoring stations (measuring O_3, NO, NO_2, RHC, SO_2, sulfate, wind speed, wind direction, and temperature). Objectives of the site selection were the following:

- Determination of the peak basin wide O_3 levels.
- Determination of RHC and NO_x concentrations in major source regions.
- Determination of characteristic flow patterns in the basin.

The closest existing air quality monitoring stations to each of the ten sites selected by the staff members are listed in table III-2. Of the stations listed, the following were used in analyzing model sensitivity to aerometric data inputs:

- Downtown Los Angeles
- Azusa
- Burbank
- Long Beach
- Anaheim
- San Bernardino
- Reseda
- Lennox
- Prado Park
- Upland

Next, each of the airshed model sensitivity simulations is identified in greater detail.

TABLE III-2. AIR QUALITY MONITORING STATIONS IN THE SOUTH COAST AIR BASIN AND THE NUMBER OF TIMES THEY WERE SELECTED FOR INCLUSION IN THE REDUCED AEROMETRIC DATA SET

Station	Number of Times Selected	Station	Number of Times Selected
Downtown Los Angeles	8	Ventura	2
Azusa	6	Ojai	0
Burbank	6	Pt. Mugu	0
West Los Angeles	1	Redlands	1
Long Beach	5	Fontana	2
Reseda	6	Upland	3
Pomona	2	Chino	1
Lennox	6	Hemet	0
Whittier	2	Banning	0
Newhall	3	Perris	1
Lancaster	0	Anaheim	5
Pasadena	1	Costa Mesa	1
Lynwood	0	Laguna Beach	0
El Toro	0	La Habra	0
San Juan Capistrano	0	Redondo Beach	0
Los Alamitos	2	Camarillo	2
Thousand Oaks	0	San Bernardino	4
Simi Valley	0	Riverside	1
Santa Paula	2	Prado Park	4
Port Hueneme	0	El Monte	3

C. ATTRIBUTES OF EACH SENSITIVITY RUN

1. Simulations Focusing on a Limited Number of Aerometric Monitoring Stations

a. Run 1--1974 Upper Air Meteorological Data

This simulation is intended to address the situation in which upper air meteorological data acquisition in a city is limited to routine, twice daily vertical soundings, customarily performed at airports. Upper air winds (about 300 meters above mean sea level) were based on the twice daily soundings (0700 and 1300 PST) at El Monte. Mixing depths for the entire region were based on these two soundings and on the hourly averaged temperatures at all available surface wind and air quality stations.

b. Run 2--1975 Upper Air Meteorological Data

This run is the counterpart of run 1; it focuses on the 4 August 1975 meteorological regime.

c. Run 3--1974 Surface and Upper Air Meteorological Data

This simulation is intended to address the situation in which both surface and upper air data acquisition in an urban area is sparse. Meteorological fields were constructed using data available from the 10 surface monitors identified earlier and from the 0700 and 1330 PST El Monte soundings.

d. Run 4--1975 Surface and Upper Air Meteorological Data

This simulation is the counterpart of run 3; it focuses on the 4 August 1975 meteorological regime.

e. Run 5.1--Ground Level Initial and Boundary Conditions

For this situation, initial and boundary conditions for 26 June 1974 were prepared from concentration data available from the reduced set of 10 aerometric stations. Meteorological fields were those of the June base case.

f. *Run 5.2--1974 Reduced Meteorological and Air Quality Data*

The purpose of this simulation was to examine the impact on model predictions of limited data from both meteorological and air quality monitoring networks. Meteorological fields from simulation 3 were used in conjunction with the initial and boundary conditions prepared from air quality data from stations nearest to the 10 sites identified in run 5.1. The 26 June episode was used for this simulation.

2. Simulations Focusing on More Specialized Aerometric Monitoring Activities

In establishing the June and August base cases, a detailed examination of ambient air quality data was carried out. Among the data resources utilized in preparing initial and boundary conditions for the base cases were the following:

> Monitoring data from the ARB station located atop Mt. Lee.

> ARB early morning hydrocarbon speciation measurements at various stations during the summers of 1974 and 1975.

> Air quality observations at monitoring stations located upwind and outside of the modeling region.

> Special studies data collected in the course of airborne monitoring over the SCAB (for example see Husar et al., 1977; Blumenthal, White, and Smith, 1978; and Edinger, 1973).

These information sources, together with the routine data gathered on the days of interest, were of great value in preparing inputs for the base cases. The existence of all of the above information was not known, however, at the time the first base case runs were made. It was only through subsequent diagnostic analyses and further investigation that certain of the additional information resources became known. Thus, though some of the preliminary attempts at establishing June and August base cases were set aside based upon additional information, these simulations are of value from the viewpoint of model sensitivity. Five of these runs are presented below.

a. Run 6--Upper Air Quality Data

Run 6 was a simulation of 26 June in which the hydrocarbon and NO_x precursor concentrations in level three grid cells* were increased (relative to the base case) based upon the detailed aircraft studies reported by Husar et al. (1977). In the absence of upper air concentration data on 26 June, pollutant concentrations aloft could only be speculated.

b. Run 7--1975 Clean Air Boundary Conditions

Run 7 differed from the 4 August base case in that lower precursor concentrations aloft were considered and inflow boundary conditions near Thousand Oaks and Laguna Beach-Costa Mesa were estimated to be similar to "clean air" concentrations. In contrast, the base case run involved the use of actual hourly concentration measurements at Thousand Oaks, Santa Paula, Simi, Costa Mesa, and Laguna Beach.

c. Run 8--1974 Clean Air Boundary Conditions

The initial and boundary conditions differ from the base case in that lower hydrocarbon and nitrogen oxide concentrations were considered ("clean air" conditions).

d. Run 9--1974 Assumed RHC/NO_x Ratio

Prior to the acquisition of the detailed aerometric hydrocarbon data reported by Mayersohn et al. (1975, 1976) for the summers of 1974 and 1975, the initial conditions for the 26 June simulations were based upon an assumed hydrocarbon to NO_x ratio of 7 (Reynolds et al., 1979). Run 9 differs from the 26 June base case in that the former simulation involved "clean air boundary conditions" and did not take into account the observed empirical relationship between RHC and THC derived from Mayersohn's data.

e. Run 10--1975 Assumed RHC/NO_x Ratio

This run is the August counterpart of run 9.

---

* The vertical structure of the modeling region consisted of four layers, whose combined height was a constant 1000 m above local ground level.

3. Simulations Focusing on Details in Emissions Inventories

Altogether, nine sensitivity runs focused on details in the various components of an airshed model emission inventory.

a. Run 11--1974 Hydrocarbon Emissions Speciation

The potential need for detailed hydrocarbon speciation of the major emission sources (as compared to overall basin wide hydrocarbon splits) was addressed in Run 11. Table III-3 gives the hydrocarbon splits used in both base cases by source category.

Composite hydrocarbon splits were developed by determining the percent contribution of olefins, paraffins, aromatics, carbonyls and ethylene to the total reactive hydrocarbon inventory. The resultant breakdown between the various species groupings are as follows:

Species	Percent by Weight (as carbon)
Olefins	5
Paraffins	70
Aromatics	17
Carbonyls	5
Ethylene	3

b. Run 12--1975 Hydrocarbon Emissions Speciation

The overall hydrocarbon splitting factors in Runs 11 and 12 were identical.

c. Run 13--1974 Mobile Source

The base case motor vehicle emission inventory employed in this study was generated by the Direct Travel Impact Model (DTIM), a successor to the Caltrans Link Emissions Model (FWY011). Advantages of the DTIM code are its ability to allocate emissions due to cold starts, hot starts, and hot soaks based on the origin, destination, and type of trip. The trip assignment program accounts for intrazonal and terminal traffic volumes in addition to the link volume records treated by its predecessor, FWY011. DTIM uses hot and cold start and hot soak files for emission estimations based on the correspondence between the traffic zones and grid cells.

TABLE III-3. HYDROCARBON SPLITS BY SOURCE CATEGORY

SOURCE CATEGORY	CES DESCRIPTION	OLE	PAR	ARO (MOL/GRAM X 100)	CARB	ETH
1	EMISSION	1.57E-01	1.57E-01	0.	0.	1.13E 00
2	.MOTOR VEHICLE	1.23E-01	4.35E 00	2.80E-01	1.84E-01	1.23E-01
3	.SHIPPING	0.	0.	0.	0.	0.
4	.RAILROAD	9.23E-02	4.79E 00	5.94E-02	5.67E-01	2.31E-01
5	.INDUSTRIAL	1.78E-01	2.74E 00	4.34E-03	7.70E-01	2.36E-03
6	.PETROLEUM	0.	0.	0.	0.	0.
7	.COMMERCIAL AND INSTITUTIONAL	0.	3.51E 00	2.84E-01	4.83E-01	2.23E-01
8	.AIRCRAFT	0.	0.	0.	0.	0.
9	.AGRICULTURAL	0.	0.	0.	0.	0.
10	..PETROLEUM MARKETING	0.	2.37E-01	2.03E-04	1.72E-03	2.26E-03
11	..DEGREASING	0.	0.	0.	0.	0.
12	..PETROLEUM REFINING	0.	0.	0.	0.	0.
13	..PETROLEUM PRODUCTION	0.	0.	0.	0.	0.
14	..CATALYST GASOLINE EXHAUST	0.	0.	0.	0.	0.
15	..CHEMICAL INDUSTRY	1.25E-01	3.28E 00	4.39E-02	5.23E-01	4.33E-01
16	..SURFACE COAT AND AIR DRY ARCH	0.	0.	0.	0.	0.
17	..AGRICULTURAL CONTROL BURN	0.	0.	0.	0.	0.
18	..POWER PLANT BOILERS	0.	0.	0.	0.	0.
19	..AIRCRAFT PISTON EXHAUST	3.85E-01	3.86E 00	2.02E-01	6.00E-01	2.88E-01
20	..JET EXHAUST	1.92E-01	5.23E 00	1.90E-01	6.66E-02	0.
21	..GASOLINE CRANKCASE	0.	0.	0.	0.	0.
22	..DRY CLEANING	0.	0.	0.	0.	0.
23	..GASOLINE EVAP CARBURETOR	0.	0.	0.	0.	0.
24	..PESTICIDES	0.	3.71E 00	1.96E-01	0.	3.03E-01
25	..WOOD PROCESSING	0.	0.	0.	0.	0.
26	..DOMESTIC SOLVENT USE	0.	4.53E 00	0.	2.31E-01	0.
27	..WASTE BURNING AND WILDFIRES	0.	0.	0.	0.	0.
28	..BALLASTING	1.78E-01	2.74E 00	4.34E-03	7.70E-01	2.38E-03
29	..NONCATALYST GASOLINE EXHAUST	0.	0.	0.	0.	0.
30	..PURGING	1.78E-01	2.74E 00	4.34E-03	7.70E-01	2.38E-03
31	..MINERAL INDUSTRY	7.09E-02	1.78E 00	5.43E-04	1.47E-02	6.91E-02
32	..FOOD PROCESSING	2.35E-02	3.41E-01	3.26E-04	2.73E-02	4.90E-02
33	..OFF LOADING	0.	5.45E 00	0.	0.	1.77E-02
34	..DIESEL EXHAUST	0.	0.	0.	0.	0.
35	..METALLURGICAL INDUSTRY	0.	0.	0.	0.	0.

TABLE III-3 (Continued)

SOURCE CATEGORY	CES DESCRIPTION	OLE	PAR	ARO	CARB	ETH
				(MOL/GRAM X 100)		
36	..ORCHARD HEATING	0.	0.	0.	0.	0.
37	..DIESEL EVAPORATIVE	0.	0.	0.	0.	0.
38	.OFF ROAD MOTOR VEHICLES	1.08E-01	4.00E 00	2.98E-01	2.67E-01	1.85E-01
39	..MOVING SHIPS	0.	0.	0.	0.	0.
40	...PET STORAGE AT GAS STATIONS	1.16E-01	6.17E 00	6.42E-02	0.	2.88E-02
41	...AIR DRIED SURFACE COATING	0.	0.	0.	0.	0.
42	...HALOGEN DEGREASING	0.	0.	0.	0.	0.
43	...SYNTHETIC DRY CLEAN	0.	5.10E 00	0.	0.	0.
44	..INDUSTRIAL SURFACE COATING	0.	1.72E 00	0.	0.	7.13E-02
45	...VEH REFUEL AT GAS STATIONS	0.	0.	0.	0.	0.
46	...PERCHLOROETHYLENE DRY CLNG	0.	0.	0.	0.	0.
47	...NON HALOGEN DEGREASING	0.	1.42E-01	0.	0.	0.
48	...HEAT TREATED SURFACE COAT	0.	1.72E 00	0.	0.	7.13E-02
49	..EXT COMB BOILERS AND HEATERS	0.	0.	0.	0.	0.
50	..ROCKET EXHAUST	0.	0.	0.	0.	0.
51	..INDUSTRIAL INCINERATION	6.33E-02	4.96E-01	0.	1.21E-01	6.42E-02
52	..DOMESTIC UTILITY 4 STROKE	0.	0.	0.	0.	0.
53	..DOMESTIC UTILITY 2 STROKE	0.	0.	0.	0.	0.
54	.DOMESTIC	0.	0.	0.	0.	0.
55	...NAT GAS POW PLANT BOILER	0.	1.93E 00	2.17E-02	2.66E-01	1.10E-02
56	...RESID OIL POW PLANT BOILER	0.	2.28E 00	0.	1.80E 00	0.
57	...COAL POWER PLANT BOILER	0.	0.	0.	0.	0.
58	..INDUSTRIAL OFF-ROAD VEH	0.	0.	0.	0.	0.
59	..CONSTRUCTION OFF-ROAD VEH	9.23E-02	4.79E 00	5.94E-02	5.67E-01	2.31E-01
60	..RECREATIONAL OFF-ROAD VEH	0.	0.	0.	0.	0.
61	..FARM OFF-ROAD VEH	0.	0.	0.	0.	0.
62	..DOMESTIC FUEL COMBUSTION	0.	0.	0.	0.	0.
63	..JET FUEL EVAPORATION	0.	0.	0.	0.	0.
64	..STRUCTURAL FIRES	0.	6.25E 00	0.	0.	0.
65	STATIONARY (GENERAL)	0.	0.	0.	0.	0.
66	..WINE PROCESSING	0.	0.	0.	0.	0.
67	...DISTILLATE OIL POW PLANT	0.	3.52E 00	0.	1.62E 00	0.
68	...PROCESS GAS POW PLANT	0.	0.	0.	0.	0.
69	...IND EX COMB RESIDUAL OIL	0.	2.28E 00	0.	1.88E 00	0.
70	...IND EX COMB DISTILLATE OIL	0.	3.52E 00	0.	1.62E 00	0.

TABLE III-3 (Continued)

SOURCE CATEGORY	CES DESCRIPTION	OLE	PAR	ARO (MOL/GRAM X 100)	CARB	ETH
71	...IND EX COMB NATURAL GAS	0.	1.93E 00	2.17E-02	2.66E-01	1.19E-02
72	...IND EX COMB PROCESS GAS	4.16E-01	2.95E 00	0.	2.53E-01	0.
73	....REFIN RESIDUAL OIL	0.	2.20E 00	0.	1.88E 00	0.
74	....REFIN DISTILLATE OIL	0.	3.52E 00	0.	1.62E 00	0.
75	....REFIN NATURAL GAS	0.	1.93E 00	2.17E-02	2.66E-01	1.19E-02
76	....REFIN PROCESS GAS	4.16E-01	2.95E 00	0.	2.53E-01	0.
77	....REFIN EX COMB BOILER	0.	0.	0.	0.	0.
78	...PET PROD EX COMB BOILER	0.	2.19E 00	1.82E-02	4.84E-01	9.98E-03
79	..POWER PLANT INT COMB ENG	7.65E-03	8.67E-03	0.	2.53E 00	9.98E-03
80	....IND AIR DRIED PAINT	0.	2.13E 00	7.26E-01	3.74E-01	3.65E-02
81	..INDUSTRIAL INT COMB ENG	0.	4.73E-01	0.	6.03E-02	3.53E-02
82	..COMMERCIAL INT COMB ENG	0.	0.	0.	9.99E-01	0.
83	...PET PROD INT COMB ENGINES	1.13E-02	8.13E-02	0.	0.	2.60E-01
84	...REFIN INT COMB ENGINES	0.	4.78E-01	0.	3.33E-02	3.57E-02
85	...PRIMARY METAL INDUSTRY	5.93E-02	2.20E-01	6.19E-03	1.31E-02	6.42E-01
86	...SECONDARY METAL INDUSTRY	0.	2.89E 00	9.66E-03	5.98E-01	5.23E-03
87	...PET PROD INDUST PROCESSES	0.	2.62E 00	3.76E-02	0.	1.25E-01
88	...PET PROD FIX RF CRUDE EVAP	0.	5.45E 00	1.09E-04	0.	1.79E-02
89	...PET PROD FLOAT CRUDE EVAP	0.	5.45E 00	0.	0.	1.77E-02
90	...REFIN INDUSTRIAL PROCESS	2.84E-01	3.22E 00	2.90E-03	4.38E-01	7.9[illegible]E-03
91	...REFIN STORAGE EVAP	3.55E-03	4.01E 00	1.07E-01	2.57E-01	2.52E-02
92	....REFIN FIX ROOF GAS EVAP	1.20E-01	6.27E 00	4.53E-02	0.	2.70E-02
93	....REFIN FIX ROOF CRUDE EVAP	0.	5.31E 00	1.52E-02	0.	0.
94	....REFIN FLOAT ROOF GAS EVAP	1.20E-01	6.27E 00	4.53E-02	0.	2.70E-02
95	....REFIN FLOAT RF CRUDE EVAP	0.	5.31E 00	1.52E-02	0.	0.
96	....PETMKT FIXED ROOF GAS EVAP	1.20E-01	6.21E 00	4.53E-02	0.	2.70E-02
97	....PETMKT FIX RF CRUDE EVAP	0.	5.94E 00	0.	0.	0.
98	....PETMKT FLOAT ROOF GAS EVAP	1.20E-01	6.27E 00	4.53E-02	0.	2.70E-02
99	....PETMKT FLOAT RF CRUDE EVAP	0.	5.94E 00	0.	0.	0.
100	...PET MARKETING STORAGE EVAP	4.99E-03	4.81E 00	0.	2.86E-01	3.19E-02
101	....PETMKT LOADING GAS EVAP	[illegible]	6.27E 00	4.53E-02	0.	2.71E-02
102	....PETMKT LOADING CRUDE EVAP	0.	5.14E 00	0.	0.	0.
103	....PETMKT LOADING AND UNLOAD	0.	5.27E 00	4.68E-02	2.09E-01	7.49E-03
104	....IND AIR DRIED SURFACE COAT	0.	3.92E 00	2.94E-02	1.37E 00	0.
105	....IND AIR DRIED LACQUER	0.	4.02E 00	1.20E-01	6.91E-01	1.12E-01

TABLE III-3 (Concluded)

SOURCE CATEGORY	CES DESCRIPTION	OLE	PAR	ARO (MOL/GRAM X 100)	CARB	ETH
106	....IND AIR DRIED ENAMEL	0.	4.43E 00	9.17E-02	3.64E-01	4.51E-01
107	....IND AIR DRIED PRIMER	0.	4.82E 00	7.22E-02	4.52E-01	2.32E-01
108	....IND AIR DRIED SOLVENT	0.	3.80E 00	2.14E-01	6.84E-01	9.51E-03
109	....IND AIR DRIED ADHESIVE	0.	2.68E 00	5.51E-01	4.14E-01	1.10E-01
110	...COMM OIL-BASE SURFACE COAT	0.	4.66E 00	1.21E-01	2.49E-01	2.46E-01
111	...COMM WATER-BASE SURF COAT	0.	4.66E 00	1.21E-01	2.49E-01	2.46E-01
112	...FLEXOGRAPH PRINTING	0.	4.48E 00	5.51E-04	1.39E-01	0.
113	..PRINTING	5.13E-01	1.54E 00	0.	3.05E-01	4.42E-02
114	...GRAVURE PRINTING	0.	3.50E 00	4.10E-01	7.83E-02	3.72E-01
115	..VEGETATIVE FOREST AND CITRUS	7.34E-01	4.40E 00	0.	1.47E 00	0.
116	..ANIMAL WASTES	0.	4.07E-01	0.	5.40E-02	0.
117	..LAND FILLS	7.34E-04	4.60E-02	2.03E-03	1.47E-03	2.85E-03
118	...PETRULEUM PRODUCTION SEEPS	0.	2.59E 00	0	0.	0.
119	...NON-TANKER BOILERS	9.23E-02	4.79E 00	5.94E-02	5.67E-01	2.31E-01
120	...TANKER BOILERS	9.23E-02	4.79E 00	5.94E-02	5.67E-01	2.31E-01
121	...PLEASURE CRAFT	1.00E-01	4.00E 00	2.00E-01	2.67E-01	1.85E-01
122	..VEH FUEL TANK GASOLINE EVAP	1.54E-01	5.07E 00	2.50E-01	0.	0.
123	..MOTORCYCLE EXHAUST	0.	0.	0.	0.	0.
124	..COM EX COMB BOILER AND HEATER	0.	1.93E 00	2.17E-02	2.66E-01	1.19E-02
125	...COM RESIDUAL OIL EX COMB	0.	2.28E 00	0.	1.60E 00	0.
126	...COM DISTILLATE OIL EX COMB	0.	3.52E 00	0.	1.62E 00	0.
127	...COM NAT GAS EX COMB	0.	1.93E 00	2.17E-02	2.66E-01	1.19E-02
128	...COM PROCESS GAS EX COMB	0.	1.93E 00	2.17E-02	2.66E-01	1.19E-02
129	..AIRCRAFT PISTON EVAP	0.	0.	0.	0.	0.
130	.TRANSPORTATION	0.	0.	0.	0.	0.

Running emissions are allocated based on link volumes of traffic. The DTIM code provides estimates of up to 6 classes of RHC, CO, NO_x, SO_x, particulates, and fuel consumption. DTIM estimates emissions for eight vehicle types instead of the four vehicle types treated by FWY011. The base case motor vehicle inventory is based upon an updated (as of 1976) origin-destination survey.

For the 1974 mobile source sensitivity simulation, an outdated motor vehicle emission assessment model (FWY011) was utilized. Features of the FWY011 Link Emissions Model include the following:

- > Emissions are calculated by link.
- > Emission estimates are provided for five pollutants--RHC, THC, CO, NO_x, SO_x.
- > Four vehicle types are treated.
- > Cold starts, hot starts and hot soaks are assigned by set percentage of link volumes.
- > Intrazonal and terminal volumes are not taken into account.

The latter two features represent deficiencies in the FWY011 approach when it is compared with DTIM. Emissions due to cold starts should be assigned to the grid of origin. Emissions due to hot soaks should be assigned to zone or grid of destination. Inputs to the FWY011 code included 1967 origin-destination data projected to 1974 assumed population growth conditions.

In essence, the major differences between the two motor vehicle inventories is that the sensitivity run inventory is based on outdated origin-destination data and does not adequately treat trip end emissions.

d. Run 14--1974 Gas Sales

The objective of this sensitivity run was to examine the effect of using data related to vehicle fuel mileage to estimate daily emissions of major pollutants from on-road motor vehicles. Although these estimates should be less precise than those derived from sophisticated motor vehicle emission models, such as the DTIM, emission estimates derived from fuel mileage are usually much easier to calculate.

The estimation of motor vehicle emissions for Run 14 was based on three major parameters:

- Vehicle fuel mileage
- Vehicle emission factors
- Area fuel sales.

1) Vehicle Fuel Mileage

Fuel mileage in miles per gallon (mi/gal) was derived from the EMFAC5 computer program. EMFAC5 has been developed by the ARB to estimate motor vehicle emission factors and fuel mileage. The EMFAC5 program is similar to the EPA's MOBILE 1 program, but it includes, among other things, the motor vehicle emission standards applicable to vehicles sold in California. EMFAC5 was used to estimate average fuel mileage values for both gasoline and diesel vehicles. A mix of the following vehicle types for the year 1974 was used in the calculations:

	Vehicle Type	Percent of Total Vehicles
Gasoline		
	Light duty automobile	80.4
	Light duty truck	12.1
	Medium duty truck	1.4
	Heavy duty gasoline	2.5
	Motorcyle	1.1
Diesel		
	Heavy duty diesel	2.5

The vehicle mix percentages listed above were developed for the SCAB by Caltrans. Based on this vehicle mix and other assumptions described below for EMFAC5, average gasoline and diesel vehicles had the following fuel mileages:

- Gasoline: 12.73 mi/gal
- Diesel: 5.5 mi/gal.

2) Vehicle Emission Factors

Emission factors in grams per mile (g/mi) were derived from EMFAC5 for hydrocarbons (HC), carbon monoxide (CO), and nitrogen oxides (NO_x). In addition to the mix of vehicles identified above, the following EMFAC5 assumptions were made to calculate emission factors:

> Average vehicle speed: 19.6 miles per hour.

> Percentage of cold start, hot start, and hot stabilized operation: 21%, 27%, and 52%, respectively.

> Ambient temperature: 75° F.

The assumptions for vehicle speed and percentage of vehicle operation in different modes were based upon the Federal Test Procedure for motor vehicles. If better data in these areas had been available, it might have been possible to calculate improved motor vehicle emission estimates.

Emission factors were calculated for cold start, hot start, and hot stabilized operation as well as for crankcase, diurnal, and hot soak emissions. EMFAC5 also provided a non-methane hydrocarbon (NMHC) to total hydrocarbon (THC) ratio of 0.9496 for gasoline vehicle exhaust emissions. Using all of these assumptions, the following emission factors were calculated from EMFAC5:

	Emissions			
	THC	NMHC	CO	NO_x
Gasoline vehicles Exhaust + crankcase, g/mi	6.52	6.19	69.92	3.69
Diesel vehicles Exhaust, g/mi	4.41	--	34.08	21.29
All vehicles*				
Diurnal, g/day	9.42	--	--	--
Hot soak, g/soak	6.63	--	--	--

---

* These are emission factors, per vehicle, for the two types of evaporative HC emissions. Diurnal emissions cover evaporative HC during expansion of vapor in the fuel tank with changes in daily ambient temperature. Hot soak emissions include vapor loss from the carburetion system after completion of the vehicle trip. These HC emissions must be added to the exhaust emissions to get total motor vehicle emissions.

3) Area Fuel Sales

Estimates of the sale of motor vehicle fuel for a specific portion of a state can be difficult to calculate. Recognizing the need for such estimates, Caltrans has compiled gasoline sales data by county for various years. The county gasoline sales data have been developed primarily from vehicle registration data.

Initially, we attempted to construct fuel sales data for the SCAB. Because the SCAB contains portions of counties, and because information pertaining to the SCAB portion of county vehicle miles traveled (or vehicle registrations) was not readily available, fuel sales were estimated instead for the Los Angeles Regional Transportation Study (LARTS) area. Moreover, fuel sales were calculated for a modified LARTS area by deleting Ventura County and estimating sales for the following counties:

> Los Angeles

> Orange

> San Bernardino--LARTS portion only

> Riverside--LARTS portion only.

Caltrans also provided 1974 data indicating that 72 percent and 77 percent of county vehicle miles traveled occurred in the LARTS portion of San Bernardino and Riverside counties, respectively.

No information was available to differentiate average weekday sales from weekend sales, so average daily gasoline sales were calculated by dividing the annual data by 365. State wide information was available, however, from the California State Board of Equalization (CSBE) to seasonally adjust the sales data to June 1974. Therefore, on the basis of these data, the estimate of gallons of gasoline sold on an average June 1974 day in the modified LARTS area was 12.31 million.

County wide sales data for diesel fuel were not available for this study. However, the CSBE estimates total state wide diesel and compressed natural gas (CNG) sales on a monthly basis. No precise information was available to subtract CNG sales from these figures, but the CSBE estimated CNG sales to be less than 5 percent of the total figures (West, 1979). Thus, no adjustment was made to these sales data to account for CNG sales.

To estimate 1974 diesel fuel sales in the modified LARTS area, the following equation was used:

$$\text{1974 daily gallons of diesel in modified LARTS} =$$

$$\frac{\left(\begin{array}{c}\text{1974 annual gallons of}\\ \text{diesel in state}\end{array}\right)\left(\begin{array}{c}\text{1974 daily gallons of}\\ \text{gasoline in modified LARTS}^*\end{array}\right)}{\text{1974 annual gallons of gasoline in state}} \qquad (3\text{-}1)$$

This number was then seasonally adjusted to June 1974 using the CSBE diesel sales data described above. On the basis of these data, the estimated gallons of diesel sold on an average June 1974 day in the modified LARTS area were 1.178 million.

4) Diurnal and Hot Soak Emissions

Two additional numbers, vehicle registrations and trips, were needed to calculate HC emissions from vehicles during diurnal and hot soak conditions. Total automobile and truck registrations were calculated from 1974 county registration data compiled by Caltrans and the California Department of Motor Vehicles, for use with the diurnal emission factor. Caltrans also provided the percent of vehicles registered in the LARTS portion of San Bernardino and Riverside counties. Using these data, 6,149,152 vehicles were registered in the modified LARTS area in 1974.

Caltrans has estimated that 24,211,000 vehicle trips occurred daily in the LARTS area in 1974. Adjusting this number on the basis of vehicle registrations to account for trips occurring in Ventura County, 23,218,000 daily trips were estimated to occur in the modified LARTS area in 1974.

5) Summary of Emissions Based on Fuel Sales

On the basis of the data previously discussed, the following motor vehicle emission estimates were calculated for June 1974 in the modified LARTS area.

---

* Not seasonally adjusted.

Emissions, tons/day	
THC	1,390
NMHC	1,330
CO	12,310
NO_x	790

The grid used for this air quality modeling study is a subset of the modified LARTS area. A comparison of the motor vehicle emission estimates derived from fuel mileage for the grid area to those derived from DTIM is given below.

	Emissions, tons/day		
	RHC*	CO	NO_x
Fuel mileage technique	1,260	11,890	760
DTIM	850	8,250	720

e. Run 15--1975 Gas Sales

The motor vehicle inventories for runs 14 and 15 are identical.

f. Run 16--Point Source

The baseline emission inventory supplied by the ARB contained power plant emissions for a "typical summer week day". However, in preparing inputs for the 26 June and 4 August base case simulations, we deleted the nominal emission rates and, in their place, substituted the actual hourly emission rates estimated by the Southern California Edison Company and the Los Angeles Department of Water and Power. Power plant emissions for this sensitivity run were based on nominal annual emission rates compiled by the District (SCAPCD, 1976) and were temporally distributed according to a diurnal profile developed by General Research Corporation (GRC, 1974).

g. Run 17--1974 Area Source

The objective of this simulation was to examine the need for spatially resolved area and point source emission data. A total of 323 point sources are included in the baseline inventory. Sixty-nine of these are thermal electric power plants. These sources are identified by UTM coordinates and various stack parameters. The rest of the emission sources in the inventory, whether they are line sources, small point sources, or diffuse area sources, are allocated according to a 25 km^2 grid mesh.

---

* Reactive hydrocarbons.

In this simulation, emissions from the following sources were unchanged from the base case:

> All 323 elevated point sources

> All on-road motor vehicles

> All natural or geo-genic sources.

The emissions from the remaining sources were aggregated on an hourly basis and allocated to the grid based upon the 1975 demographic distribution.

h. Run 18--1975 Area Source

The area source emissions files for runs 17 and 18 were identical.

i. Run 19--Temporal Resolution

In this study, the temporal distribution of categories of area sources was varied to examine the change in overall emission patterns and resultant air quality modeling concentrations. For each of the source categories constituting the area source emission file, a temporal distribution of diurnal emissions was estimated from data sources described below.*

1) Temporal Distribution of Existing Emissions Inventory

The area source portion of the Los Angeles Air Quality Maintenance Plan (AQMP) emission inventory provided to SAI for use in this study had been disaggregated into categories of emission sources (CES) with assigned CES numbers. A system of codes was established to readily identify the temporal distributions used to resolve daily emissions into specific hours of the day. Four coding series were used to identify the applicable temporal distributions of emissions in the AQMP inventory.*

---

* The spatial resolution of area sources was unchanged from the base case for this run.

> 0,1,2,...22,23 refers to the individual clock hours of the day; for example, 13 means 1300 to 1400.

> Numbers in the 40 and 50 series are derived from the equation:

$$40 + n = 0700 + 100n \text{ from 0700 to the final time.} \qquad (3\text{-}2)$$

For example, 41 means 0700 to 0800 and 50 means 0700 to 1700.

> Numbers in the 70 and 80 series are derived from the equation:

$$70 + n = 0.5(100n) \text{ equally distributed around 1200.} \qquad (3\text{-}3)$$

For example, 78 means 0800 to 1600.

> 98 refers to a distribution over each of the 24 hours of the day.

In addition, each of the hours that is assigned emissions by an individual code receives an equal distribution of the total emissions.

The AQMP temporal distributions represented by the codes were determined from survey responses of sources in the Los Angeles area and from assumptions regarding typical source operations. As seen in table III-4, some source categories within the AQMP inventory had emissions distributed in only one manner (e.g., CES number 38), though other categories had emissions distributed by more than one distribution (e.g., CES numbers 121 and 19). Furthermore, many categories had as many as 12 applicable temporal distributions, presumably as a result of responses to the source survey. However, the AQMP inventory provided to SAI did not differentiate the distributions that had been assumed for a source category from those that had resulted from the source survey.

2) Temporal Distributions Developed by SAI

For each source category, an assumed temporal distribution representative of a summer season was chosen. The principle behind each choice was to select one specific distribution for a source category that was

---

* For purposes of this discussion, "0700" and "0700 to 0800" mean one hour of emissions beginning at 0700 and ending at 0800.

TABLE III-4. CATEGORIES OF EMISSION SOURCES AND TEMPORAL DISTRIBUTIONS

	CES Number	AQMP Distribution Code*	Assumed Distribution Code÷
Emission	1	98(56)	112
Transportation	130	§	
• Motor Vehicle	2	**	**
• • Catalyst gasoline exhaust	14		
• • Non-Catalyst gasoline exhaust	29		
• • Gasoline evap. loss carb	23		
• • Gasoline evap. loss fuel tank	122	47(100)	49
• • Gasoline crankcase	21		
• • Diesel exhaust	34		
• • Diesel evaporative	37		
• • Motorcycle exhaust	123		
• Off Road Motor Vehicle	38	78(100)	50
• • Industrial	58		
• • Construction	59	78(100)	78
• • Recreational	60		
• • Farm	61		
• Shipping	3		
• • Purging	30		
• • Off loading	33		
• • Ballasting	28	98(100)	98
• • Transit	39		
• • • Boilers non-tankers	119	98(100)	98
• • • Boilers tankers	120	98(100)	98
• • • Pleasure craft	121	84(67),78(29)	82
• Railroad	4	98(100)	98
• Aircraft	8		
• • Jet exhaust	20	**	**
• • Jet fuel evaporation	63		
• • Piston exhaust	19	51(85)	51
• • Piston fuel evaporation	129		
• • Rocket	50		
Stationary	65		
• Petroleum	6		
• • Production	13		
• • • Ext. combustion boilers	78	98(82)	98

TABLE III-4 (continued)

	CES Number	AQMP Distribution Code*	Assumed Distribution Code†
· · · Int. combustion engines	83	98(100)	98
· · · Industrial processes	87	98(96)	98
· · · Seeps	118	98(100)	98
· · · Crude oil evap. fixed roof	88	98(99)	98
· · · Crude oil evap. floating roof	89	98(100)	98
· · Refining	12		
· · · Ext. combustion boilers	77		
· · · · Boilers residual oil	73	98(100)	98
· · · · Boilers distillate oil	74	98(94)	98
· · · · Boilers natural gas	75	98(83)	98
· · · · Boilers process gas	76	98(100)	98
· · · Internal combustion engines	84	98(100)	98
· · · Industrial processes	90	98(75)	98
· · · Storage evap.	91	98(94)	98
· · · · Crude oil evap. fixed roof	93	98(100)	98
· · · · Crude oil evap. floating roof	95	98(100)	98
· · · · Gasoline evap. fixed roof	92	98(94)	98
· · · · Gasoline evap. floating roof	94	98(100)	98
· · Marketing	10	98(99)	98
· · · Storage evap.	100	98(64)	98
· · · · Crude oil evap. fixed roof	97	98(100)	98
· · · · Crude oil evap. floating roof	99	98(100)	98
· · · · Gasoline evap. fixed roof	96	98(83)	98
· · · · Gasoline evap. floating roof	98	98(100)	98
· · · Loading and Unloading	103	98(49),78(16)	112
· · · · Gasoline evap.	101	98(59),78(26)	112
· · · · Crude oil	102	98(62)	112
· · · Underground storage @ stations	40	52(99)	52
· · · Vehicle refueling @ stations	45		
· Commercial & Institutional	7	98(100)	112
· · Internal combustion engines	82	98(100)	98
· · Ext. combusion boilers & space heat	124	98(100)	98
· · · Residual oil	125	98(82)	98
· · · Distillate oil	126	98(70)	98
· · · Natural gas	127	98(78)	98
· · · Process gas	128	98(100)	98
· · Printing	113	98(23),78(23),56(23)	112
· · · Flexigraphic	112	78(26),98(18),56(13)	112
· · · Gravure	114	98(42),78(17),54(17)	112
· · Surface coating air dried achit.	16		
· · · Oil base including solvent	110	78(100)	50
· · · Water base	111	78(100)	50
· · Dry Cleaning	22		
· · · Petroleum base perchlorethylene	46	78(22),49(12),47(11),44(10)	50
· · · Synthetic	43	78(23),47(15),46(11),44(11)	50
· · Degreasing	11	78(44),56(22)	101
· · · Halogenated	42	78(26),56(22)	101
· · · Non-Halogenated	47	78(23),56(12)	101

TABLE III-4 (continued)

	CES Number	AQMP Distribution Code*	Assumed Distribution Code
. Industrial	5	98(35),78(20)	112
. . Internal combustion engines	81	98(71)	112
. . External combustion boilers & heaters	49		
. . . Residual oil	69	98(64)	112
. . . Distillate oil	70	98(60)	112
. . . Natural gas	71	98(66)	112
. . . Process gas	72		
. . Chemical	15	98(36),56(12),78(10)	98
. . Metallurgical	35		
. . . Primary Metals	85	98(53)	112
. . . Secondary Metals	86	98(62)	112
. . Mineral	31	78(34),98(23)	112
. . Wood Processing	25		
. . Elec. generation boiler	18		
. . . Residual oil	56		
. . . Distillate oil	67		
. . . Natural gas	55		
. . . Process gas	68		
. . . Coal	57		
. . Elec. generation Inter. Comb.	79	98(56)	98
. . Surface coating	44	98(60),44(40)	101
. . . Heat treated	48	78(27)	101
. . . Air dried	41		
. . . . Paint	80	44(18),78(15),43(14),42(12)	101
. . . . Varnish and Shellac	104	78(25),98(18),46(11),44(11)	101
. . . . Lacquer	105	78(28),44(14)	101
. . . . Enamel	106	78(28),44(12),42(10)	101
. . . . Primer	107	78(35),56(16)	101
. . . . Solvent	108	78(43),98(15)	101
. . . . Adhesives	109	78(26),56(14)	101
. . Incineration	51	43(18),42(18),41(14),78(11)	52
. . Land fills	117	98(100)	98
. Agricultural	9		
. . Agricultural control burn	17		
. . Vegetative forest & citrus	115	98(100)	98
. . Animal wastes	116	98(100)	98
. . Pesticides	24	52(100)	50
. . Food processing	32	98(38),78(19)	112
. . Orchard heating	36		
. . Waste burning or wildfires	27		
. . Wine processing	66		
. Domestic	54		
. . Solvent use	26	52(100)	134
. . Utility equipment 2 stroke	53		

TABLE III-4 (concluded)

	CES Number	AQMP Distribution Code*	Assumed Distribution Code+
. . Utility equipment 4 stroke	52	98(100)	6,7,8...21,22
. . Fuel combustion	62		
. . Structural fires	64		

---

* Temporal distribution codes provided with the emissions inventory. Numbers in parentheses represent the percent of the total number of grid cells that have emissions (for that source category) temporally distributed by that code; only those codes and their corresponding percentages most responsible for distributing the category's emissions are listed. See text for explanation of codes.

+ Temporal distribution codes assumed by SAI. See text for explanation of codes.

§ No emissions associated with that CES number were provided in this portion of the emissions inventory.

** The temporal distribution of emissions provided with the emission inventory was used for the assumed distribution. For CES #2, each of the 24 hours received almost equal weighting (i.e., about 4 percent). For CES #20, each of the 24 hours received some weighting; from 0700 through 2200 inclusive, each hour received almost equal weighting (i.e., about 6 percent).

typical of the normal operating schedules of that source type. In other words, an individual who was familiar with typical source operations, but was lacking site-specific operating data, would be expected to chose temporal distributions similar to those chosen by SAI for this study. For some source categories, no change in source operation was assumed (e.g., CES number 59). However, the majority of sources were assumed to have a single temporal distribution different from the distributions used in the AQMP inventory (e.g., CES numbers 121 and 19).

Three new temporal distributions were developed to account for the complex nature of many source operations. These distributions were used to accommodate the differing operating schedules of individual sources within a category. The three new distributions and their codes, which were developed by SAI, refer to the following hours and weighting of those hours.

> 101 refers to a distribution of hours from 0700 to 2200; each hour from 0700 to 1700 receives twice the weighting of the hours from 1700 to 2200. Thus,
>
> - 0700 to 1700: each hour receives 8 percent of the emissions.
>
> - 1700 to 2200: each hour receives 4 percent of the emissions.
>
> - All other hours receive no emissions.

> 112 refers to a distribution over each of the 24 hours of the day; each hour from 0800 to 1600 receives twice the weighting of the hours from 1600 to 0800. Thus,
>
> - 0800 to 1600 : each hour receives 6.25 percent of the emissions.
>
> - 1600 to 0800: each hour receives 3.125 percent of the emissions.

> 134 refers to a distribution of hours from 0700 to 2100; each hour from 0700 to 1800 receives twice the weighting of the hours from 1800 to 2100. Thus,
>
> - 0700 to 1800: each hour receives 8 percent of the emissions.

- 1800 to 2100: each hour receives 4 percent of the emissions.

- All other hours receive no emissions.

The distributions chosen for each source category were based on four primary sources of information:

> The temporal distributions used in the AQMP emission inventory provided to SAI were partially developed from responses to a survey of sources in the Los Angeles area. From this inventory, as indicated in table III-4, SAI calculated the percent of the total number of grid cells that had emissions. These were calculated for each source category that was temporally distributed by a particular code. This information was useful in estimating the portion of a category's sources that operated on a particular schedule.

> Temporal distributions that were developed by the ARB for a study in the Sacramento, California area provided additional information on common operating schedules for many categories of sources (Reynolds et al., 1979).

> The Regional Air Pollution Study (RAPS) (Littman, 1978) provided temporal distributions for selected categories of area sources.

> Area source distributions used during the development of the Sacramento AQMP were also useful for comparison purposes. These temporal distributions were derived from survey responses and estimates of typical source operation (Skelton et al., 1977).

In addition, SAI has been developing a comprehensive emission inventory for a modeling study in central California. Temporal distributions developed in conjunction with that study also assisted this effort.

4. Simulations Focusing on Model Grid Mesh Configuration

Three simulations were carried out to explore the impact of model configuration on predicted oxidant levels; these are discussed briefly.

a. Run 20--10 km

For this simulation, all gridded model inputs were averaged to yield 10 km spatial resolution instead of the 5 km resolution of the base case.

b. Run 21--2-Layer Model

In this run, the vertical resolution of the model consisted of two layers separated by a temporally and spatially varying inversion base.

c. Run 22--1-Layer Model

In this run, the vertical resolution of the model consisted of one single layer below the inversion base.

D. CONCLUDING REMARKS

This chapter identifies the 22 model sensitivity simulations investigated in the present study. Most of these simulations correspond to those originally identified in our preliminary project work plan. Some do not, however. In certain instances (e.g., simulations involving 2.5 km grid resolution or motor vehicle emissions based on inadequate transportation models), simulations were not carried out either as a result of a directive from the EPA Project Officer or because compilation of the requisite input files was too expensive.

Some simulations discussed in this report were not originally contemplated. They arose from efforts to establish suitable June and August base cases. Previous experience in developing acceptable base case results in several urban areas--Los Angeles, Denver, St. Louis, Sacramento, and Las Vegas--indicates that an iterative process of input preparation, model simulation, and diagnostic analysis is generally necessary. When simulation results fall short of the model performance goals established in the evaluation effort, additional analyses are performed to ascertain the sources of inadequate performance. Often, information is discovered that was not initially considered in preparing model inputs. Thus, the base case simulation is generally preceded by several model runs. These intermediate simulations may be regarded then, in some sense, as based on a lesser level of detail in input information. Consequently, a few of the runs performed in the model evaluation effort have been included in the sensitivity analysis presented in later chapters.

IV BASE CASE SIMULATIONS

A. IDENTIFICATION OF ANALYSIS PROCEDURES

Several model performance measures and graphical procedures have been developed to display and interpret the airshed model simulation results (Hayes, 1978; Hillyer et al. 1979). In this section analytical procedures that are most informative in evaluating model performance in simulating the base case results are identified.

Because of the vast amount of output information available from a grid model simulation, the performance of the model may be evaluated from a variety of perspectives. Hayes (1978) reported on a detailed examination of candidate model performance measures for air quality dispersion models. Five attributes of desirable model performance were identified.

> Accuracy of the calculated peak concentration

> Absence of systematic bias

> Lack of gross error

> Temporal correlation

> Spatial alignment.

The accuracy of the calculated peak concentration can be evaluated in different ways. The observed peak concentration can be compared with the highest calculated value at a specified monitoring station, or the calculated peak value can be taken as the highest value calculated at any one of the monitoring sites, or even as the highest value in any ground-level grid cell. The basis for comparing observed and computed peak concentrations must therefore be specified. Here the focus is on the correspondence between the peak computed concentrations and observed concentrations at each monitoring station.

Absence of systematic bias means that a model does not consistently underestimate or overestimate pollutant concentrations. The presence of systematic bias can be inferred qualitatively by plotting pairs of calculated and observed concentrations on a scattergram plot. If the locus of the prediction-observation pairs falls along a 45° line (the so-called perfect correlation line), the absence of systematic bias in model calculations is indicated. If the locus of the points falls above or below the line, a systematic bias toward underestimation or overestimation is suggested. Obviously, it is quite possible for a model to exhibit a bias toward overestimation in a particular concentration range and toward underestimation in a different range. The estimates of systematic bias are calculated in the following manner.

CALCULATION OF SYSTEMATIC BIAS:

$$\text{Mean deviation} = \frac{1}{N} \sum_{i=1}^{N} (C_{c,i} - C_{m,i}) \quad , \qquad (4\text{-}1a)$$

$$\text{Mean normalized deviation} = \frac{1}{N} \sum_{i=1}^{N} \frac{(C_{c,i} - C_{m,i})}{C_{m,i}} \quad , \qquad (4\text{-}1b)$$

where C_c and C_m are the computed and measured concentrations, respectively, and N is the total number of comparisons.

Continuing with the scattergram concept, the absence of gross error can be determined by the "dispersion" of the prediction-observation pairs about the perfect correlation line. If the mean absolute deviation of the pairs about the perfect correlation line is small, the model is said to exhibit "skill" as a predictor. If the mean absolute deviation is large, the model suffers from the presence of large gross errors. Both the mean signed deviation (a measure of systematic bias) and the mean absolute deviation (a measure of gross error) can be determined as a function of concentration level. These measures can be presented either as quantities normalized by the observed concentration level or as nonnormalized values. The mean absolute deviation and the mean normalized absolute deviation are given by

CALCULATION OF GROSS ERROR:

$$\text{Mean absolute deviation} = \frac{1}{N} \sum_{i=1}^{N} |C_{c,i} - C_{m,i}| \quad , \qquad (4\text{-}2a)$$

$$\text{Mean normalized absolute deviation} = \frac{1}{N} \sum_{i=1}^{N} \frac{|C_{c,i} - C_{m,i}|}{C_{m,i}} \quad . \qquad (4\text{-}2b)$$

Temporal correlation refers to the "timing" or "phasing" of the observed and computed ozone levels at a specified station. The temporal correlation for a given station can be determined by using the pairs of hourly observed and calculated concentrations to define daily mean values. A correlation coefficient can then be calculated according to routine statistics. Lack of temporal correlation can be ascribed to one or more causes, including inadequate characterization of emissions, wind, or mixing depth inputs.

To calculate an average temporal correlation coefficient, ρ , we perform the following change of variable (Hoel, 1962):

$$\phi_j = \frac{1}{2} \ln \frac{1 + r_j}{1 - r_j} \quad , \tag{4-3}$$

where r_j is the computed correlation coefficient for Station j on the basis of hourly pairs of predictions and observations. Next, the mean value of the ϕ_j's is estimated from

$$\bar{\phi} = \frac{1}{M} \sum_{j=1}^{M} \phi_j \quad , \tag{4-4}$$

where M is the number of monitoring stations. Since the values of ϕ_j will be approximately normally distributed, it can be shown that

$$\bar{\phi} = \frac{1}{2} \ln \frac{1 + \rho}{1 - \rho} \quad , \tag{4-5}$$

where ρ is the average value of the temporal correlation coefficient. Thus, ρ can be determined from the following equation.

TEMPORAL CORRELATION COEFFICIENT:

$$\rho = \frac{\exp(2\bar{\phi}) - 1}{\exp(2\bar{\phi}) + 1} \quad . \tag{4-6}$$

The spatial alignment of observed and calculated concentration fields is another useful measure of model performance. For a given hour, imagine two concentration isopleths, one constructed from observed pollutant concentrations and the other from the corresponding model calculations. If one isopleth were placed over the other, the degree of spatial misalignment would be easy to discern (at least qualitatively). Spatial alignment can be quantified by considering a sequence of "time slices." For instance, for a particular hour, mean values of calculated and observed concentrations can be computed from the ensemble of monitoring stations. Spatial correlation coefficients can then be computed for each hour according to routine statistics. Common sources of spatial misalignment include discrepancies between the modeled and observed wind velocities, inaccuracies in the emission inventory, and the treatment of photochemistry. Estimation of the average spatial correlation coefficient follows the procedure described above.

B. BASE CASE SIMULATION RESULTS

1. Accuracy of Computed Peak Concentrations

The simplest comparison that can be made between computed and observed ozone concentrations involves the maximum hourly average values at each monitoring station. In table IV-1, the peak one-hour-average ozone predictions are compared with the highest observed concentrations for the 4 August 1975 and 26 June 1974 Los Angeles simulations. The times at which the peaks occur do not necessarily coincide. Included in the table are the percentage differences between computed and observed values for each of 17 monitoring stations. The percentage difference varies from -53 percent to +500 percent. As noted in the table, the absolute concentration level of the prediction-observation pair should be considered when examining percentage differences. Percentage differences are useful measurements at high concentration levels; absolute differences are more relevant at low concentration levels.

A more useful comparison of the peak concentration predictions can be made by considering only those values above a particular level. For example, in table IV-2 the average percentage differences between peak computed and observed concentrations are presented for:

> All stations reporting peak ozone concentrations greater than or equal to 2 pphm.

> All stations reporting peak ozone concentrations greater than or equal to 12 pphm (the National Ambient Air Quality Standard).

TABLE IV-1. SUMMARY OF MAXIMUM COMPUTED AND OBSERVED O_3 CONCENTRATIONS

Monitoring Station	26 June 1974				4 August 1975			
	Computed (pphm)	Observed (pphm)	Difference (pphm)	Difference (percent)	Computed (pphm)	Observed (pphm)	Difference (pphm)	Difference (percent)
Downtown	14	11	3	27	9	13	4	-31
Azusa	35	27	8	30	19	25	6	-24
Burbank	22	19	3	16	20	19	1	5
West Los Angeles	8	8	0	0	5	7	2	-29
Long Beach	7	5	2	40	3	4	1	-25
Reseda	10	16	4	-38	12	22	10	-45
Pomona	33	29	4	14	26	29	3	-10
Lennox	6	1	5	500	3	4	1	-25
Whittier	16	21	5	-24	7	15	8	-53
Pasadena	29	26	3	12	18	23	5	-22
Lynwood	11	6	5	83	5	4	1	25
Fontana	20	34	14	-41	20	31	11	-35
Upland	34	25	9	36	25	31	6	-19
Chino	19	27	8	-30	21	27	6	-22
Prado Park	15	--	--	--	20	23	3	-13
El Toro	10	19	9	-47	9	--	--	---
Los Alamitos	8	10	2	-20	3	--	--	---
Simi	9	13	4	-31	12	16	4	-25
Redlands	21	29	8	-28	19	28	9	-32
San Bernardino	20	26	6	-23	22	27	5	-19
Mt. Lee	17	27	10	-37	--	--	--	--

NOTES: A dash indicates missing data. The percentage difference is defined as:

$$\frac{(\text{Predicted}) - (\text{Observed})}{(\text{Observed})} \times 100\%.$$

TABLE IV-2. SUMMARY OF MODEL PERFORMANCE IN COMPUTING PEAK OZONE CONCENTRATIONS

(Average absolute percentage difference between computed and observed values*)

	Oxidant Episode	
Basis for Average Value	26 June 1974	4 August 1975
All stations $>$ 2 pphm	31	26
All stations reporting peak concentrations $>$ 12 pphm	29	25
All stations reporting peak concentrations $>$ 20 pphm	27	24

---

* The values reported in this table are calculated as follows:

$$\text{Average Absolute Percentage Difference} = \frac{1}{N} \sum_{i=1}^{N} \left| C_{c,i} - C_{H,i} \right|$$

> All stations reporting ozone concentrations greater than or equal to 20 pphm.

For both simulations, the accuracy of peak predicted ozone levels increases slightly with increasing concentration level. For ozone concentrations at or above the federal standard, peak predicted concentrations are accurate to within 25 to 30 percent of the observed levels.

An evaluation of the four other generic model performance measures discussed in the preceding section--systematic bias, gross error, temporal correlation, and spatial alignment--is presented in table IV-3. In the following sections, we provide preliminary interpretation of these results.

2. Estimates of Systematic Bias

Measures of potential systematic bias can be calculated as either nonnormalized or normalized quantities. The latter are normalized by the observed concentration level. In table IV-3, measures of bias are presented for each simulation for conditions when the observed ozone concentrations equal or exceed the 10 and 20 pphm levels. The nonnormalized bias is estimated by calculating the average (signed) difference between pairs of computed and observed concentrations (computed minus observed). The normalized bias is estimated from equation (4-1b). An appraisal of systematic bias as a function of measured concentration level can be made from figure IV-1, in which the results of the 26 June and 4 August simulations are presented. In this figure we plot the mean normalized deviations. The following conclusions can be drawn from the results.

> Computed ozone concentrations for 26 June exhibit tendencies toward overestimation at low concentration levels (between 2 and 7 pphm) and underestimation at high concentration levels ($\geq$ 20 pphm).

> Computed ozone levels for 4 August generally exhibit a tendency toward underestimation throughout the entire range of observed concentrations.

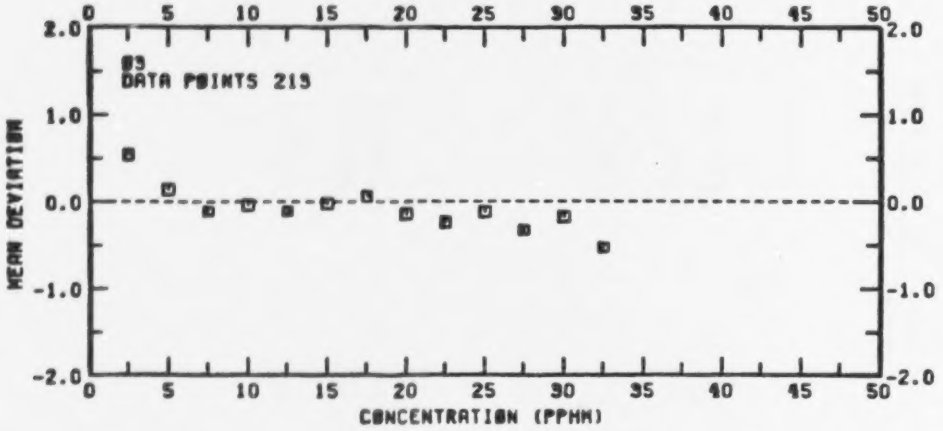
> Computed ozone concentrations exhibit greater mean deviation for 4 August than for 26 June.

TABLE IV-3. SUMMARY OF AIRSHED MODEL PERFORMANCE MEASURES FOR OZONE FOR THE LOS ANGELES SIMULATION

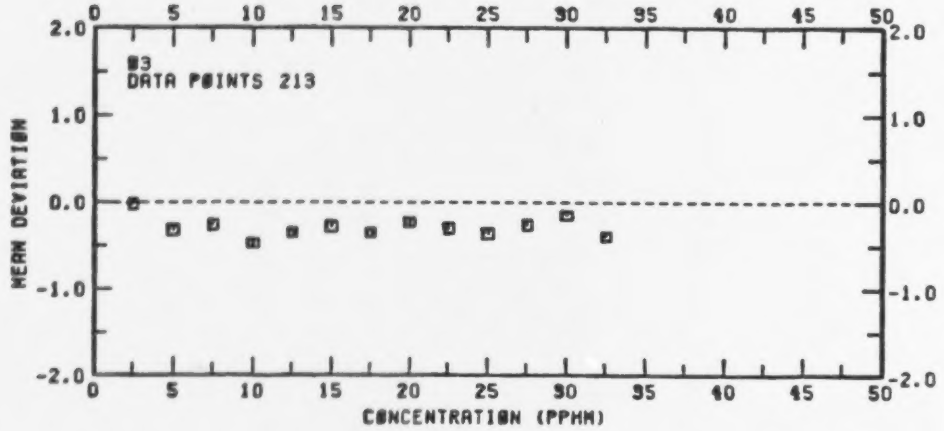
Performance Attribute	Performance Measure	26 June 1974		4 August 1974	
		O_3 > 20 pphm	O_3 > 10 pphm	O_3 > 20 pphm	O_3 > 10 pphm
Absence of systematic bias	Mean deviation				
	> Normalized-average	-.235	-.060	-.305	-.336
	> Std. dev.	.330	.403	.264	.346
	> Nonnormalized-average	-5.845 pphm	-1.502 pphm	-7.196 pphm	-5.296 pphm
	> Std. dev.	8.351 pphm	7.108 pphm	5.980 pphm	5.438 pphm
Lack of gross error	Mean absolute deviation				
	> Normalized-average	.352	.344	.322	.404
	> Std. dev.	.192	.216	.242	.263
	> Nonnormalized-average	8.706 pphm	5.867 pphm	7.536 pphm	6.268 pphm
	> Std. dev.	5.142 pphm	4.246 pphm	5.533 pphm	4.271 pphm

TABLE IV-3 (Concluded)

Performance Attribute	Performance Measure	26 June 1974		4 August 1974	
		$O_3 > 20$ pphm	$O_3 > 10$ pphm	$O_3 > 20$ pphm	$O_3 > 10$ pphm
Temporal correlation	Temporal correlation coefficients				
	> Each station	-.664 to .751	-.661 to .871	.508 to .791	.181 to .859
	> All-station average	-.171	.423	.710	.686
Spatial alignment	Spatial correlation coefficients				
	> Each hour	-.424 to .625	.028 to .657	.088 to .875	-.158 to .884
	> All-hour average	.185	.380	.414	.565



(a) 26 June 1974



(b) 4 August 1975

FIGURE IV-1. ESTIMATES OF SYSTEMATIC BIAS IN COMPUTED OZONE CONCENTRATIONS AS A FUNCTION OF OBSERVED CONCENTRATION LEVEL

3. Estimates of Gross Error

The second generic model performance measure summarized in table IV-3 is the mean absolute deviation. This indication of gross error is estimated by averaging the absolute (unsigned) difference between the pairs of computed and observed concentrations. In table IV-3 these measures are presented as both normalized and nonnormalized quantities. figure IV-2 presents the mean normalized absolute deviation as a function of measured ozone concentration level for the two simulation days. For the August base case the mean deviation generally diminishes with increasing concentration level. Typically, at low concentrations (5 to 10 pphm) the discrepancies are about 50 to 60 percent. However, near the peak concentration levels (about 25 pphm or higher) the discrepancies are reduced to roughly 35 to 40 percent.

In contrast, the 26 June results suggest a trend toward greater gross error at higher concentration levels. Above 25 pphm, the errors increase from about 25 percent to 60 percent at 33 pphm. This increase in gross error is primarily the result of large point sources of NO_x located directly upwind of three of the ozone monitors that recorded high ozone concentration levels. Titration of ambient ozone in the model by direct emission of large quantities of NO_x leads to lower predicted concentrations than if the plumes from these sources were not immediately diluted into the grid volume upon emission.[*]

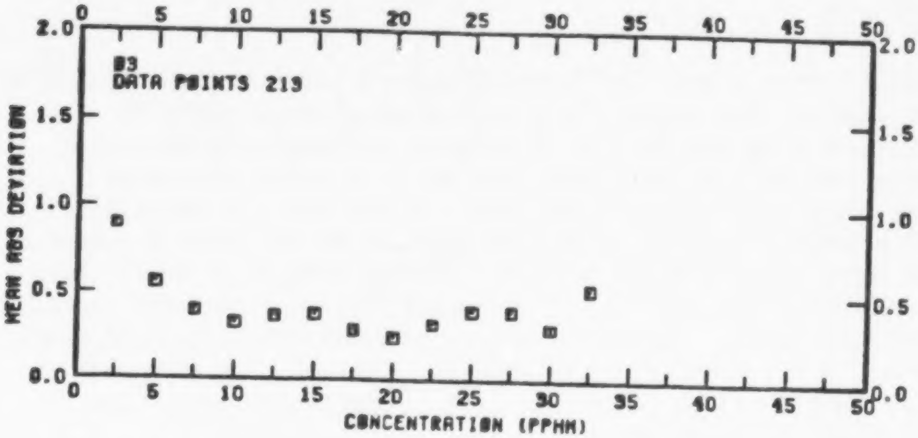
The two preceding performance measures--estimates of systematic bias and of gross error--are useful in examining the extent of model bias and the accuracies that exist for various observed concentrations. However, the simulation results can also be viewed from an overall perspective by considering differences between predictions and observations without regard to concentration level. This perspective is achieved by plotting the distribution of residuals--that is, the computed minus observed concentrations. These distributions are given in figure IV-3.

Two attributes of overall model performance can be estimated from the distribution of residuals:

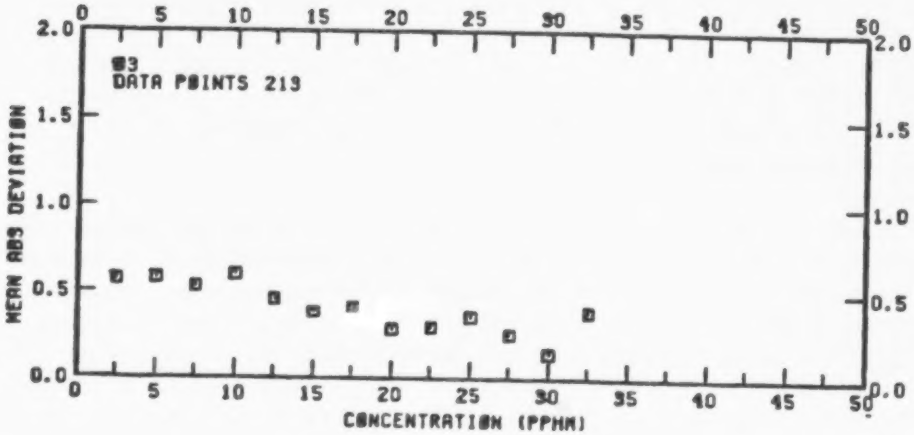
> Accuracy is the degree of conformity of a particular model prediction to an observed value (a surrogate for the true

---

[*] Work is presently underway, under sponsorship by the Electric Power Research Institute, to eliminate this problem through incorporation of a sub-grid-scale reactive plume model in the airshed model.

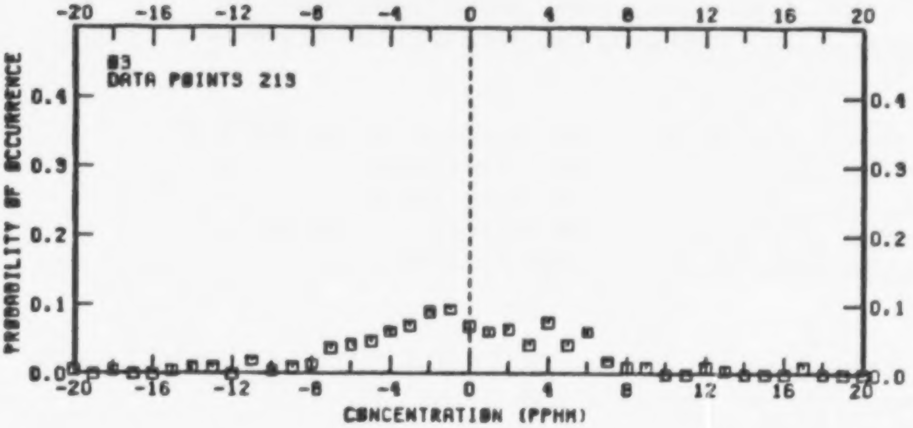


(a) 26 June 1974

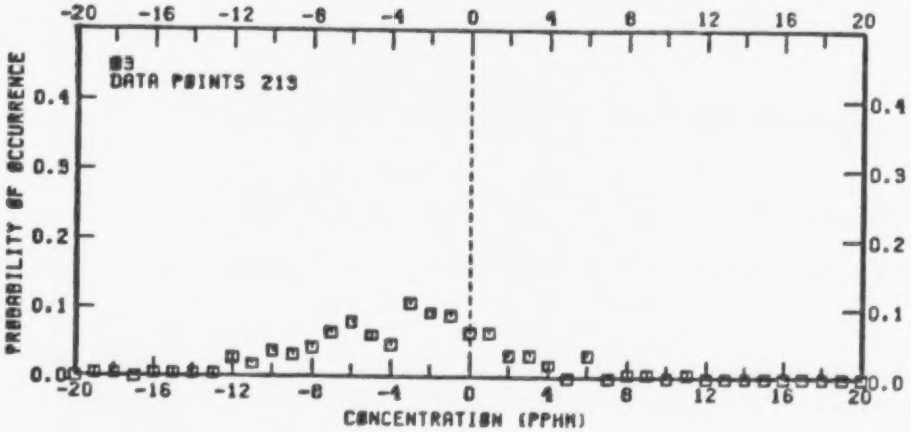


(b) 4 August 1975

FIGURE IV-2. ESTIMATES OF ERROR IN COMPUTED OZONE CONCENTRATIONS AS A FUNCTION OF MEASURED CONCENTRATION LEVEL



(a) 26 June 1974



(b) 4 August 1975

FIGURE IV-3. DISTRIBUTION OF RESIDUALS (PREDICTIONS MINUS OBSERVATIONS) FOR THE OZONE SIMULATION RESULTS

concentration). The mean value of the frequency distribution of residuals has been chosen to represent the degree of accuracy demonstrated in a particular simulation.

- Precision is the degree of conformity of the ensemble of prediction and observation pairs to a specified residual value. The standard deviation about the mean value of the difference distribution has been chosen to represent the degree of precision demonstrated in a particular simulation.

Overall accuracy of the ozone predictions for the two simulation days can be estimated from the mean (or first moment) of the difference distributions in figure IV-3. The standard deviation about the mean (the second moment) is an estimate of model precision as defined above and is also obtainable from the distributions. Analysis of the difference distribution yields the following estimates of model accuracy and precision:

	O_3 (pphm)	
Simulation Day	Accuracy	Precision
26 June 1974		± 5 pphm
4 August 1975	-3 pphm	± 5 pphm

On the basis of these results, one can conclude that the overall accuracy of the model in predicting O_3 levels was greater for 26 June than for 4 August.

4. Temporal Correlation

The temporal correlation coefficients for ozone in table IV-3 indicate a rather broad range of values for the individual monitoring stations. At concentration levels exceeding 10 pphm, average values for the coefficients are 0.423 and 0.686 for 26 June and 4 August, respectively. Of the two airshed simulations, the 4 August 1975 results exhibit the better temporal correlation. These results indicate that the "timing" of the ozone buildup and subsequent decay, when viewed over all stations, is rather variable. Peak model prediction at some stations are early, whereas at others they are late (when compared with the observed maximum). Low temporal correlation coeficients are also due to a more rapid reduction in predicted ozone levels than in observed values.

5. Spatial Alignment

Examination of the spatial correlation coefficients in table IV-3 indicates that these values are generally smaller than the corresponding temporal coefficients. For instance, at ozone levels greater than or equal to 10 pphm, the spatial coefficients are 0.380 (26 June) and 0.565 (4 August). Moreover, the spatial correlation decreases with increasing concentration level, suggesting that the location of the peak predicted "pollutant cloud" is displaced spatially from that indicated by the monitoring network. The 4 August 1975 results also demonstrate the better spatial correlation.

An additional spatial measure of model performance* is the "distance distribution." To explore further this aspect of the results, every monitoring station was examined for each hour between 1000 and 1700 PST to determine the distance from the grid cell in which the monitor is located to the nearest grid cell at which a prediction either equaled or "bracketed" the observed value. This "distribution of distances" is presented in figure IV-4. In over half of the cases a prediction comparable to a given measurement can be located within two to three grid cells--a distance of 10 to 15 kilometers. Such discrepancies can be caused by relatively small errors in wind speed or direction input data.

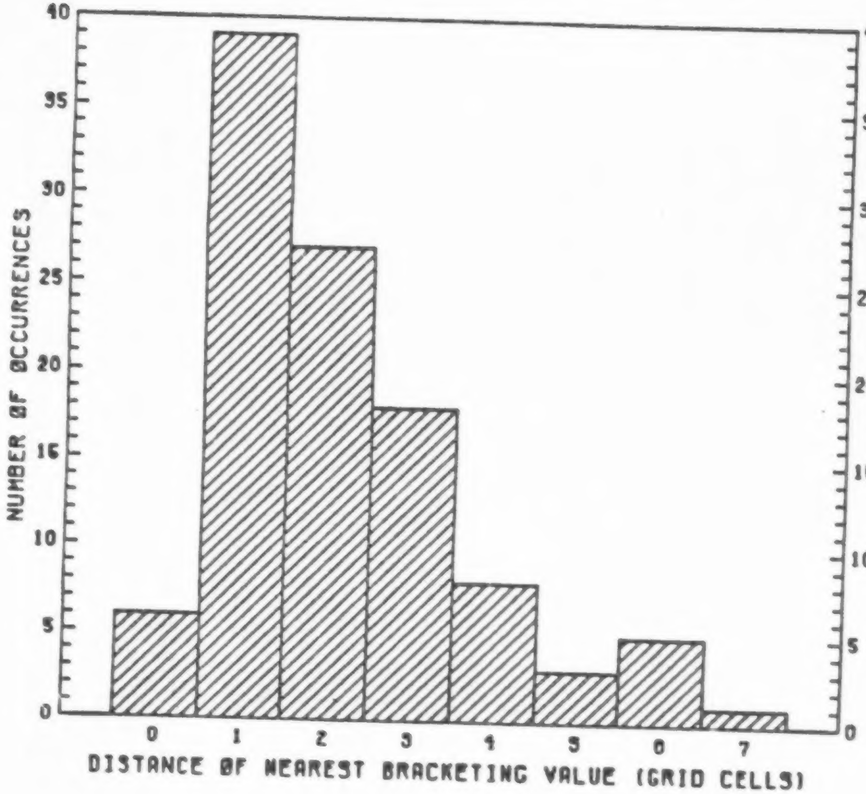
Having discussed the base case simulation results in a rather general fashion, attention is now focused on model performance at the various monitoring stations.

C. SIMULATION RESULTS FOR SPECIFIC MONITORING STATIONS

An informative though often time-consuming method for analyzing model performance is the evaluation of temporal trends in computed and observed values at the various monitoring station locations. From this procedure one can develop a conceptual picture of the formation and transport of

---

* Distance distributions, as discussed here, are not intended to be used to determine whether model performance may be judged satisfactory or unsatisfactory. Instead, this procedure is used to provide possible further insight into the possibility of biases in certain data input (e.g., wind fields), the existence of steep concentration gradients, and so forth.

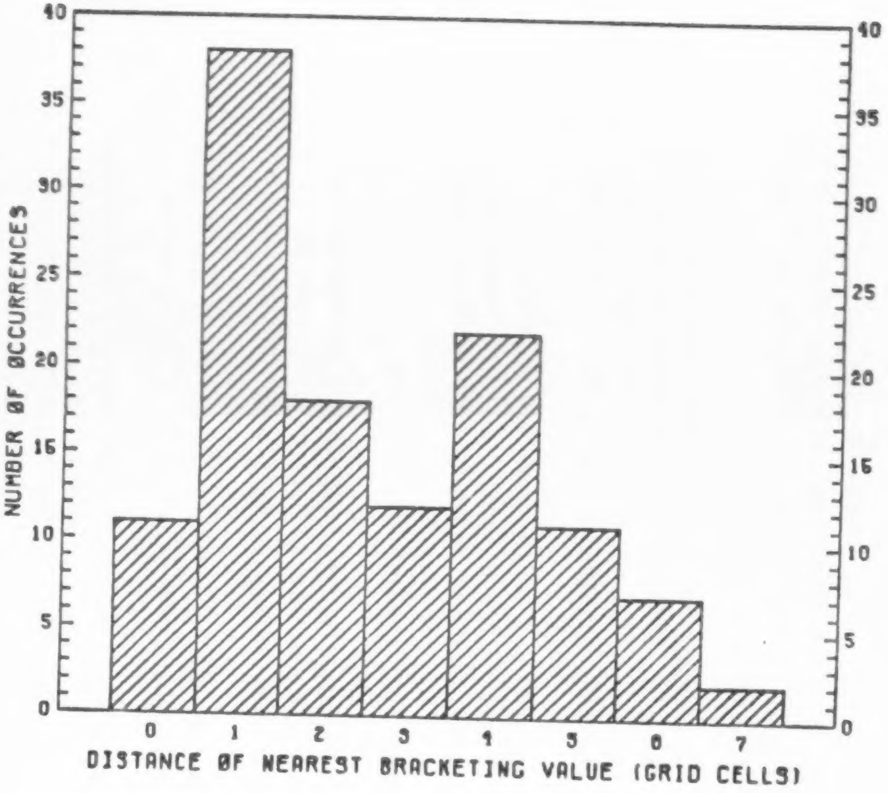


```
NO. OF VALUES MISSING             54
NO. OF VALUES BELOW THRESHHOLD 0
TOTAL VALUES ABOVE THRESHHOLD  120
NO. OF VALUES BRACKETED          107
NO. OF VALUES NOT BRACKETED     13

MEAN   = 2.2 GRID CELLS
STD DEV= 1.5 GRID CELLS
(STATISTICS BASED ON VALUES BRACKETED)
```

(a) 26 June 1974

FIGURE IV-4. NUMBER OF GRID CELLS FOR WHICH MODEL PREDICTIONS BRACKET OBSERVED OZONE CONCENTRATIONS FOR 18 STATIONS DURING THE PERIOD 10:00 a.m. TO 5:00 p.m. PST



```
NØ. ØF VALUES MISSING           42
NØ. ØF VALUES BELØW THRESHHØLD 0
TØTAL VALUES ABØVE THRESHHØLD  192
NØ. ØF VALUES BRACKETED         121
NØ. ØF VALUES NØT BRACKETED     11

MEAN    = 2.8 GRID CELLS
STD DEV= 1.8 GRID CELLS
(STATISTICS BASED ØN VALUES BRACKETED)
```

(b) 4 August 1975

FIGURE IV-4 (Concluded)

secondary pollutants, such as ozone, throughout the simulation period. The locations of the various monitoring stations and certain relevant topographical features of the SCAB are identified in figure IV-5. Figures IV-6 and IV-7 present the ozone results for the 26 June and 4 August simulations, respectively.

Upon review of the plots of computed and measured concentrations shown in figures IV-6 and IV-7, general and specific comments can be made with regard to each of the two simulations. (During the following discussion, the reader may wish to refer back to figure IV-5, which relates the locations of the monitoring stations to general features of the South Coast Air Basin.) The following comments can be made on review of the base case simulation results.

> For the 26 June 1974 simulation:

- In the San Fernando Valley predictions generally agree with observations; computed peak concentrations fall between 15 and 40 percent of the observed values.

- In the San Gabriel Valley (Pasadena, Pomona, and Azusa) a slight (10 to 30 percent) overestimation occurs; predicted levels increase more rapidly than the observed concentration levels.

- In the San Bernardino Valley (Fontana, San Bernardino, Redlands) there is a "quenching" of ozone followed by a secondary ozone buildup after noon. The model underestimates the peak concentrations by 20 to 40 percent.

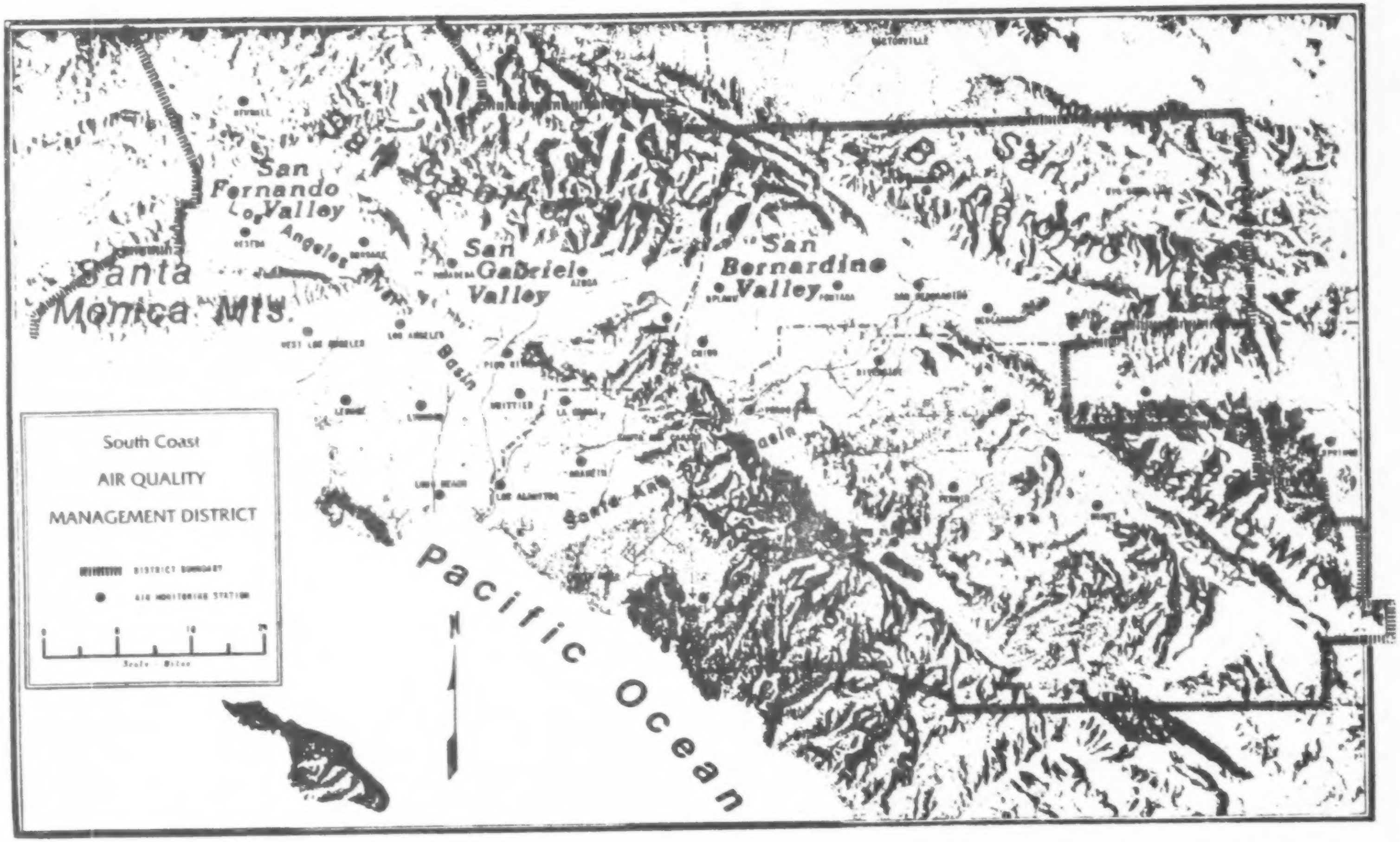
- Along the coast (Long Beach, West Los Angeles, Los Alamitos) predicted ozone levels build up sooner and diminish faster than the observed concentration levels.

> For the 4 August 1975 simulation:

- In the San Fernando Valley peak levels at Reseda are underestimated by nearly 50 percent; the peak at Burbank is overestimated by only 5 percent.

- In the San Gabriel Valley (Pasadena, Azusa) ozone is underestimated by 20 to 25 percent.

- In the San Bernardino Valley (Fontana, San Bernardino, Redlands) ozone peaks are underestimated by 25 to 35 percent. However, only at the Redlands station does the anomalous ozone quenching effect appear.



Santa Monica Mts.
San Fernando Valley
Los Angeles River
San Gabriel Mts
San Gabriel Valley
Los Angeles Basin
San Bernardino Valley
San Bernardino Mts
Santa Ana
Santa Ana Mts
San Jacinto Mts
Pacific Ocean
VICTORVILLE
WEST LOS ANGELES
LOS ANGELES
PASADENA
AZUSA
WHITTIER
LA HABRA
LONG BEACH
LOS ALAMITOS
ANAHEIM
CHINO
UPLAND
FONTANA
SAN BERNARDINO
REDLANDS
RIVERSIDE
PERRIS
HEMET
South Coast
AIR QUALITY
MANAGEMENT DISTRICT
DISTRICT BOUNDARY
AIR MONITORING STATION
0
10
Scale - Miles
N

FIGURE IV-5. TOPOGRAPHICAL FEATURES AND THE LOCATIONS OF AIR MONITORING STATIONS IN THE SOUTH COAST AIR BASIN

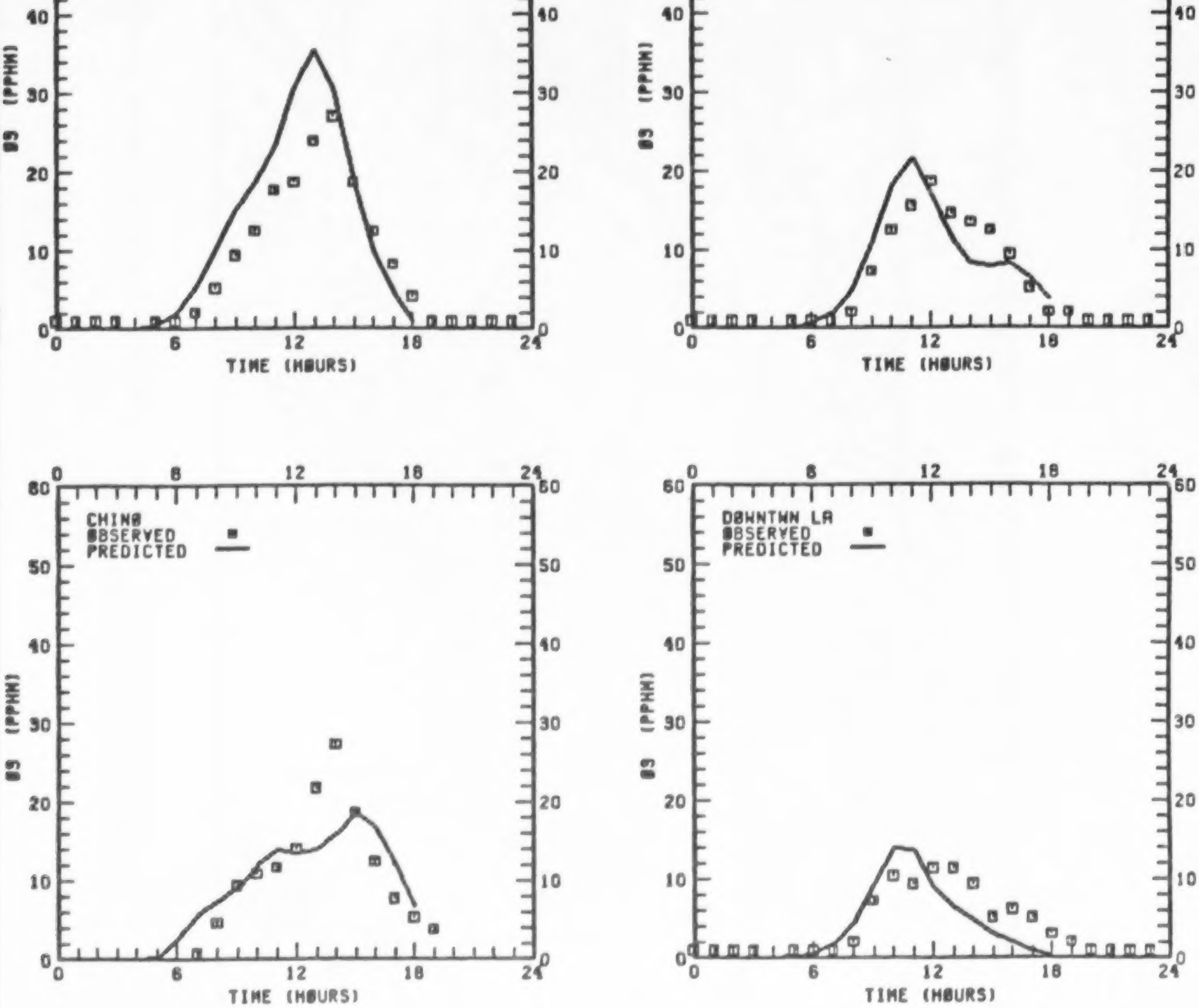


FIGURE IV-6. CALCULATED AND OBSERVED OZONE CONCENTRATIONS FOR 26 JUNE 1974

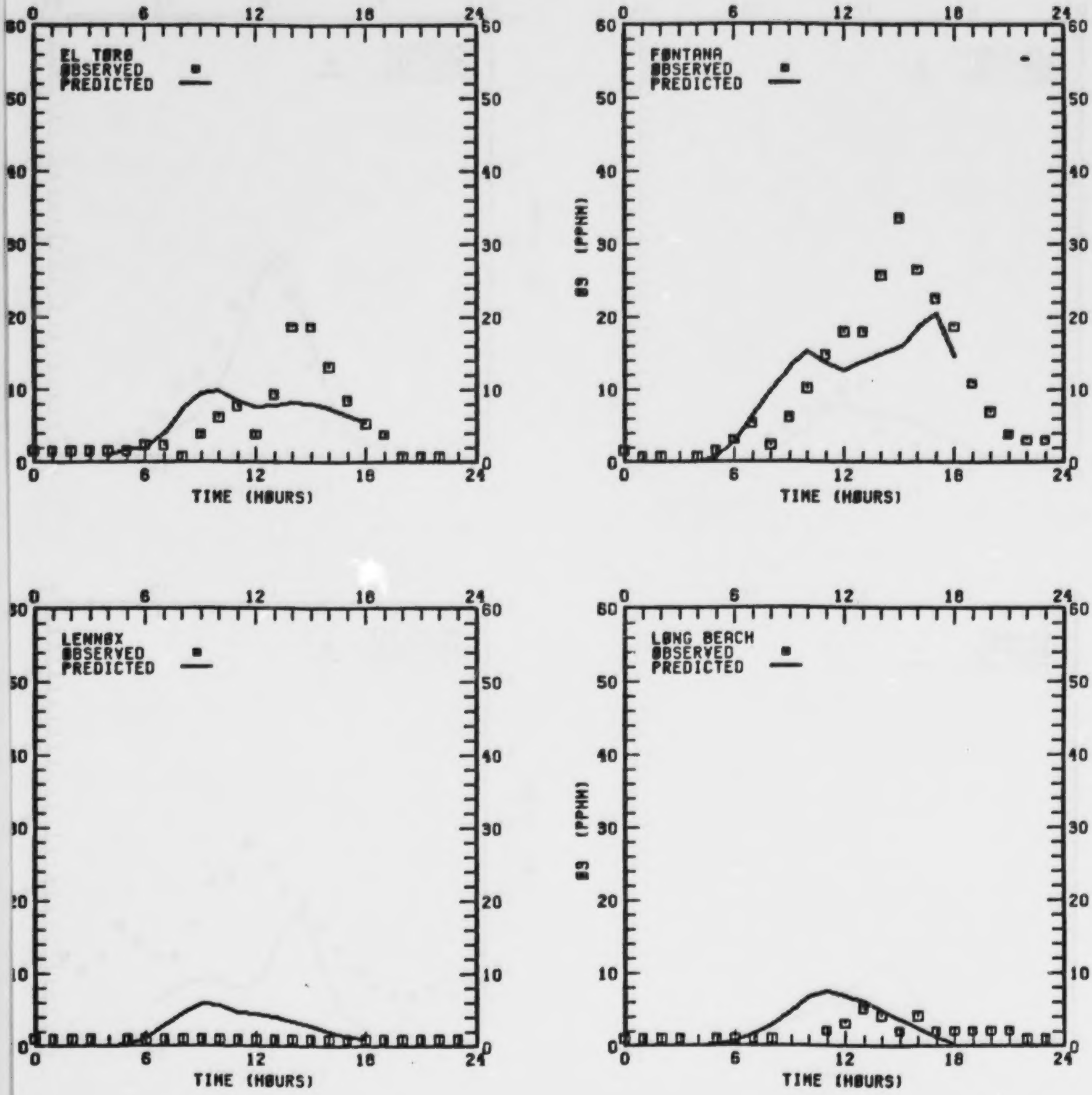


FIGURE IV-6 (Continued)

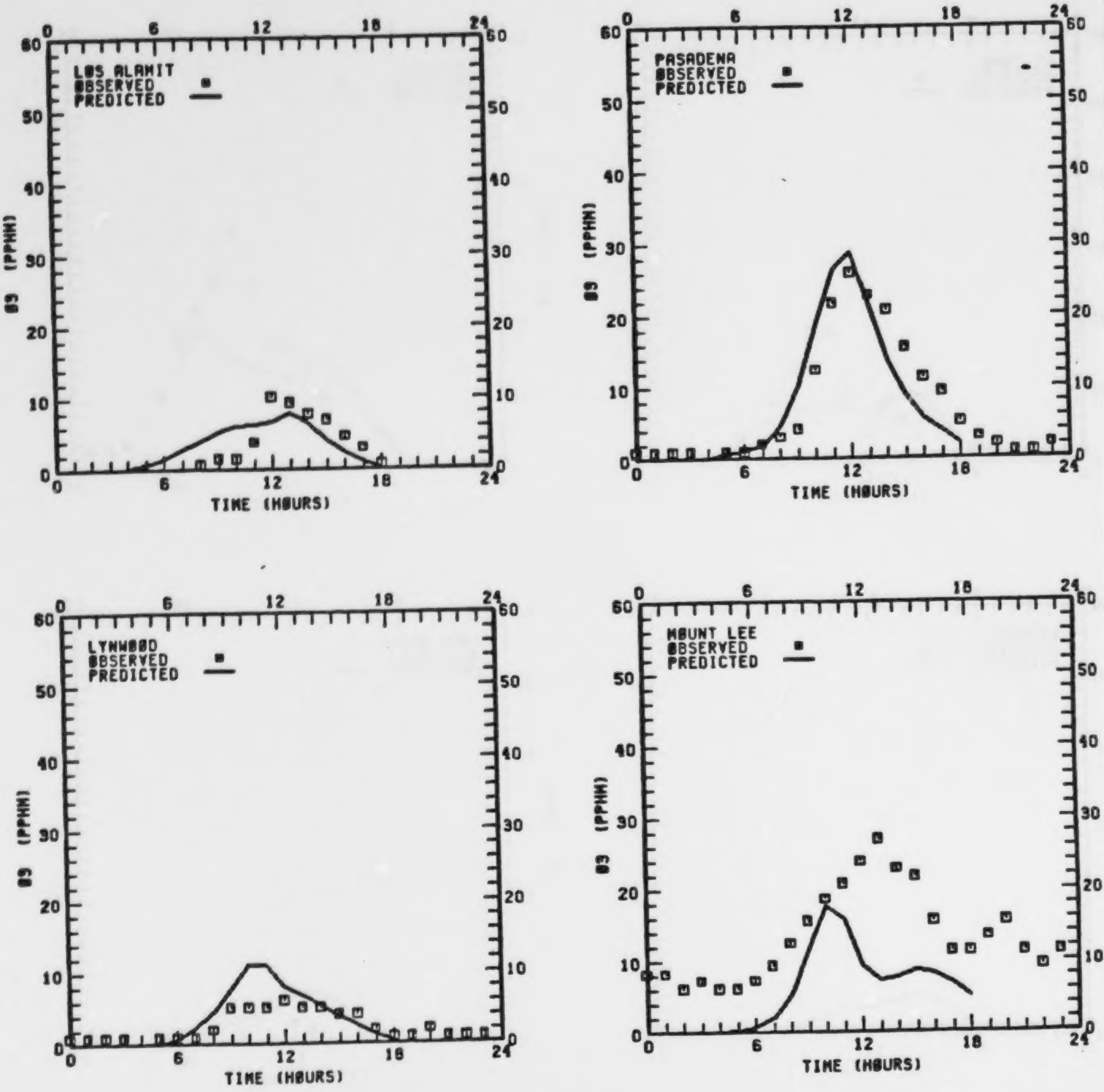


FIGURE IV-6 (Continued)

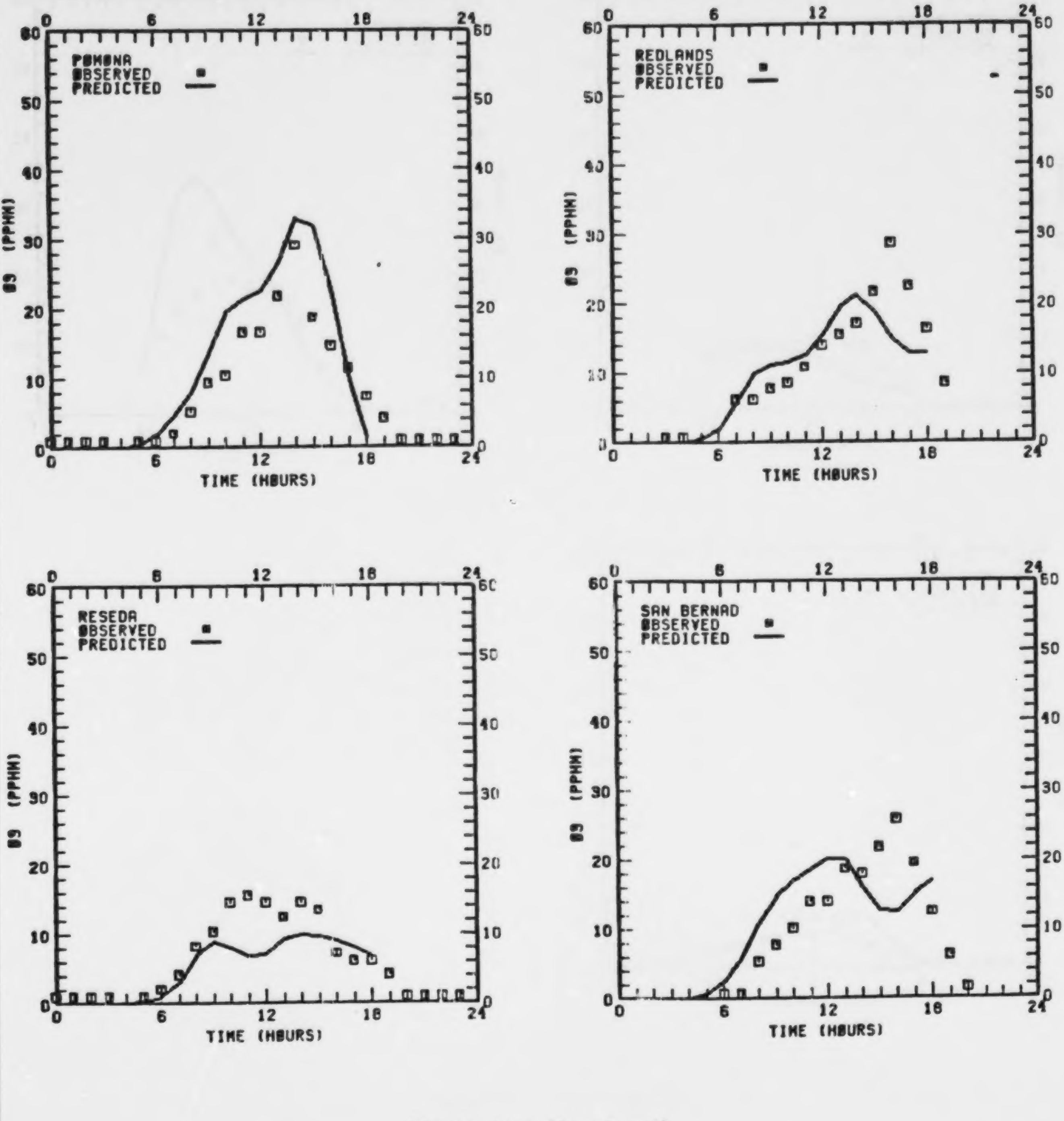


FIGURE IV-6 (Continued)

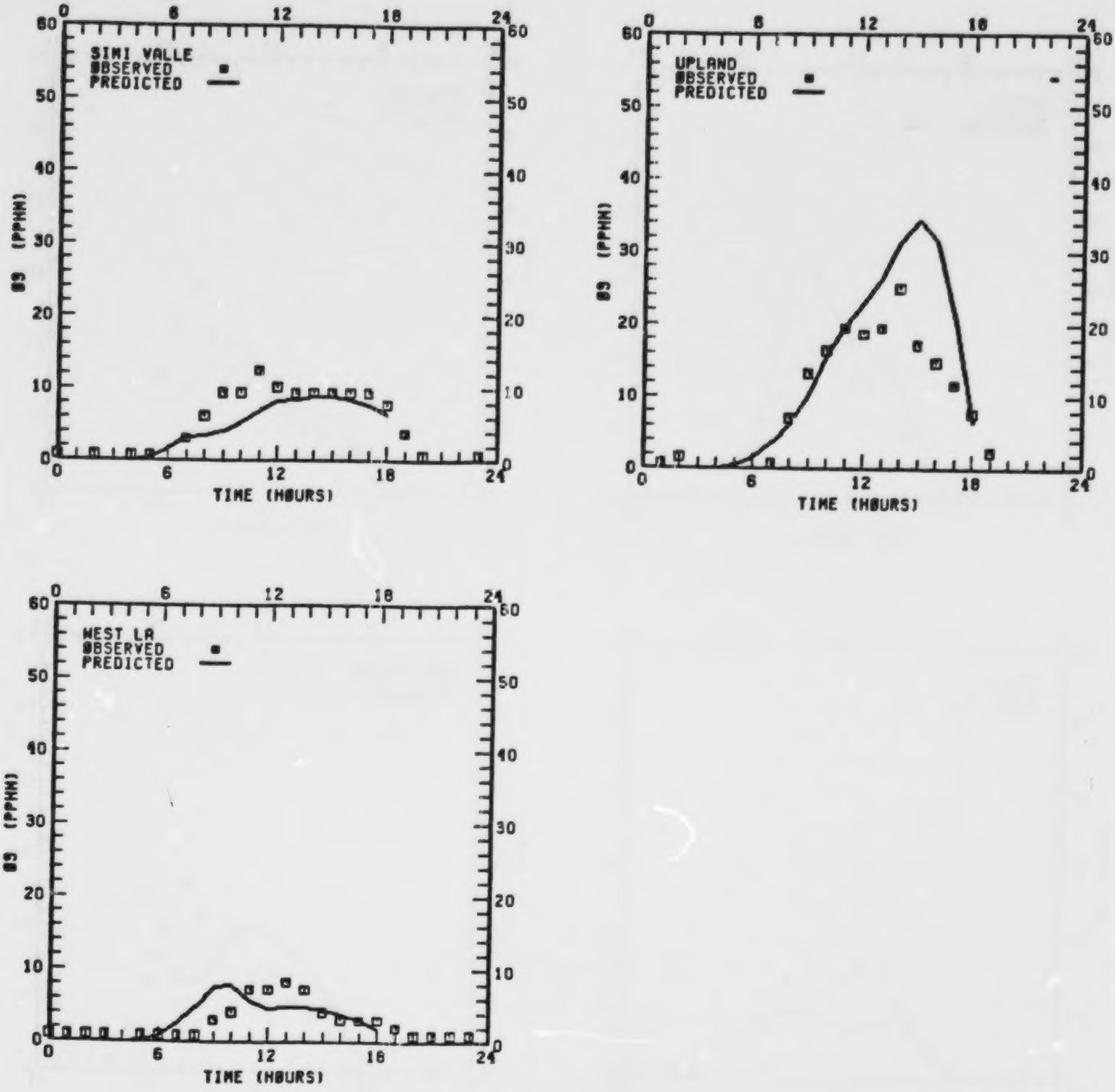


FIGURE IV-6 (Concluded)

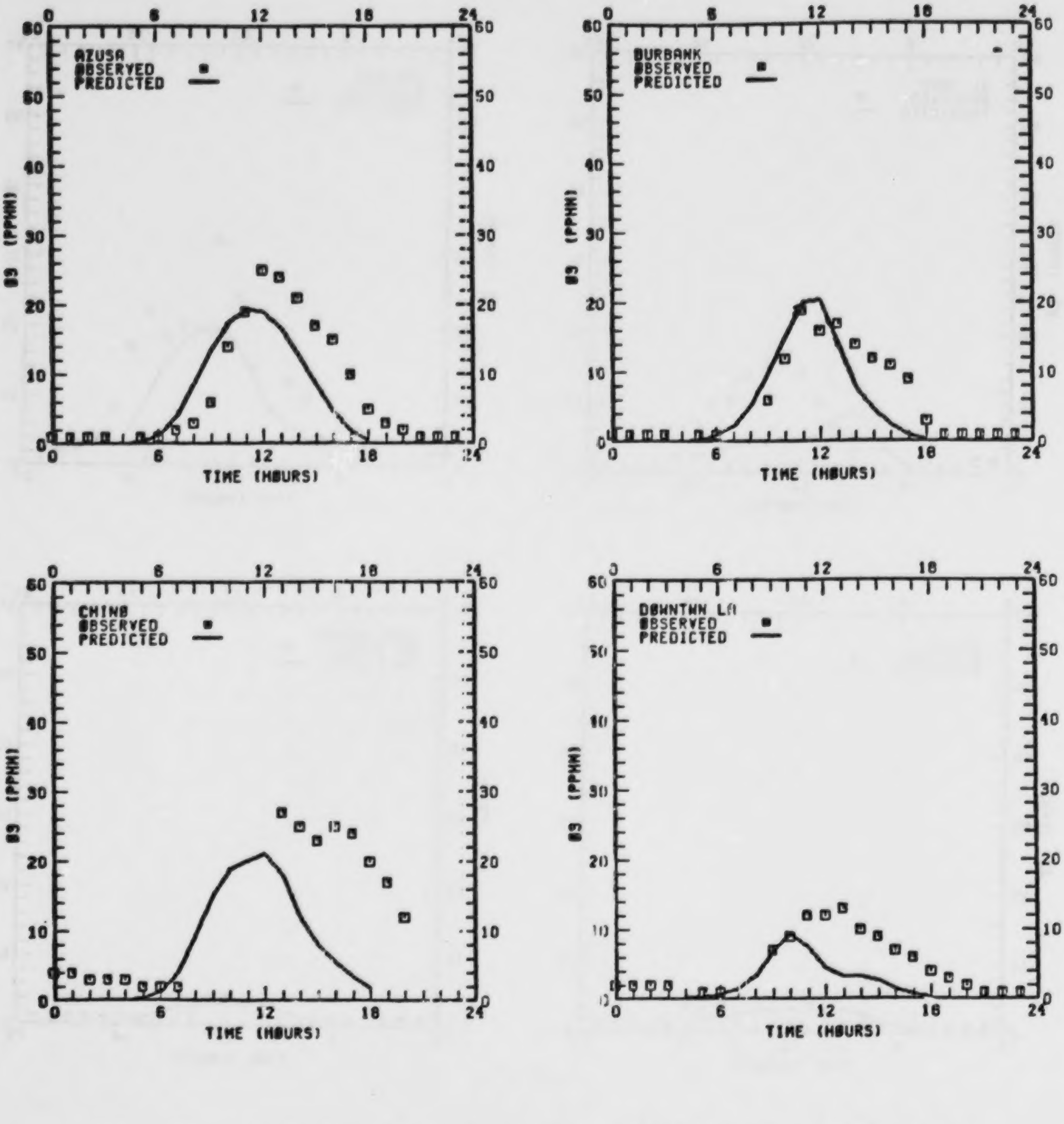


FIGURE IV-7. CALCULATED AND OBSERVED OZONE CONCENTRATIONS FOR 4 AUGUST 1975

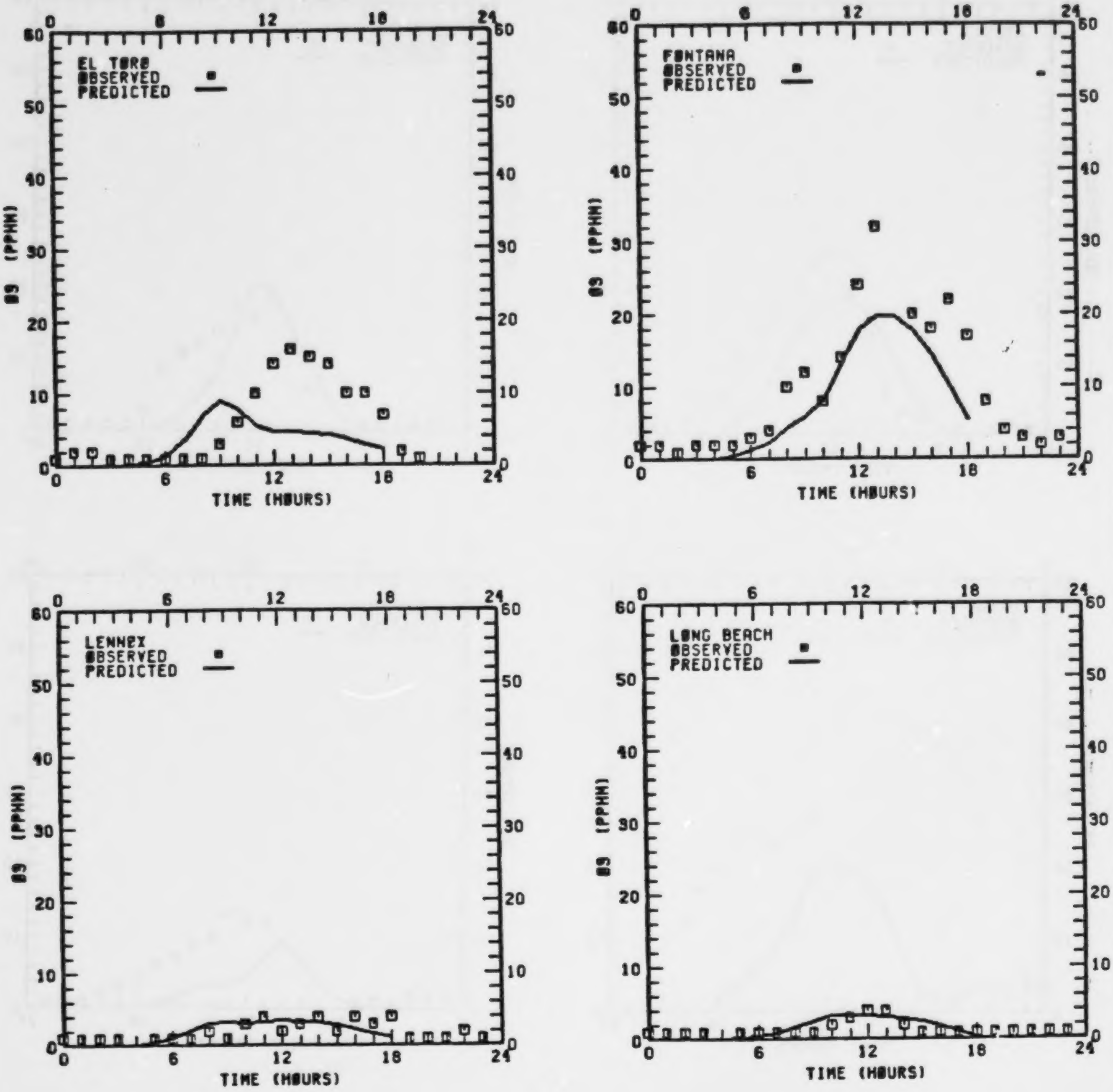


FIGURE IV-7 (Continued)

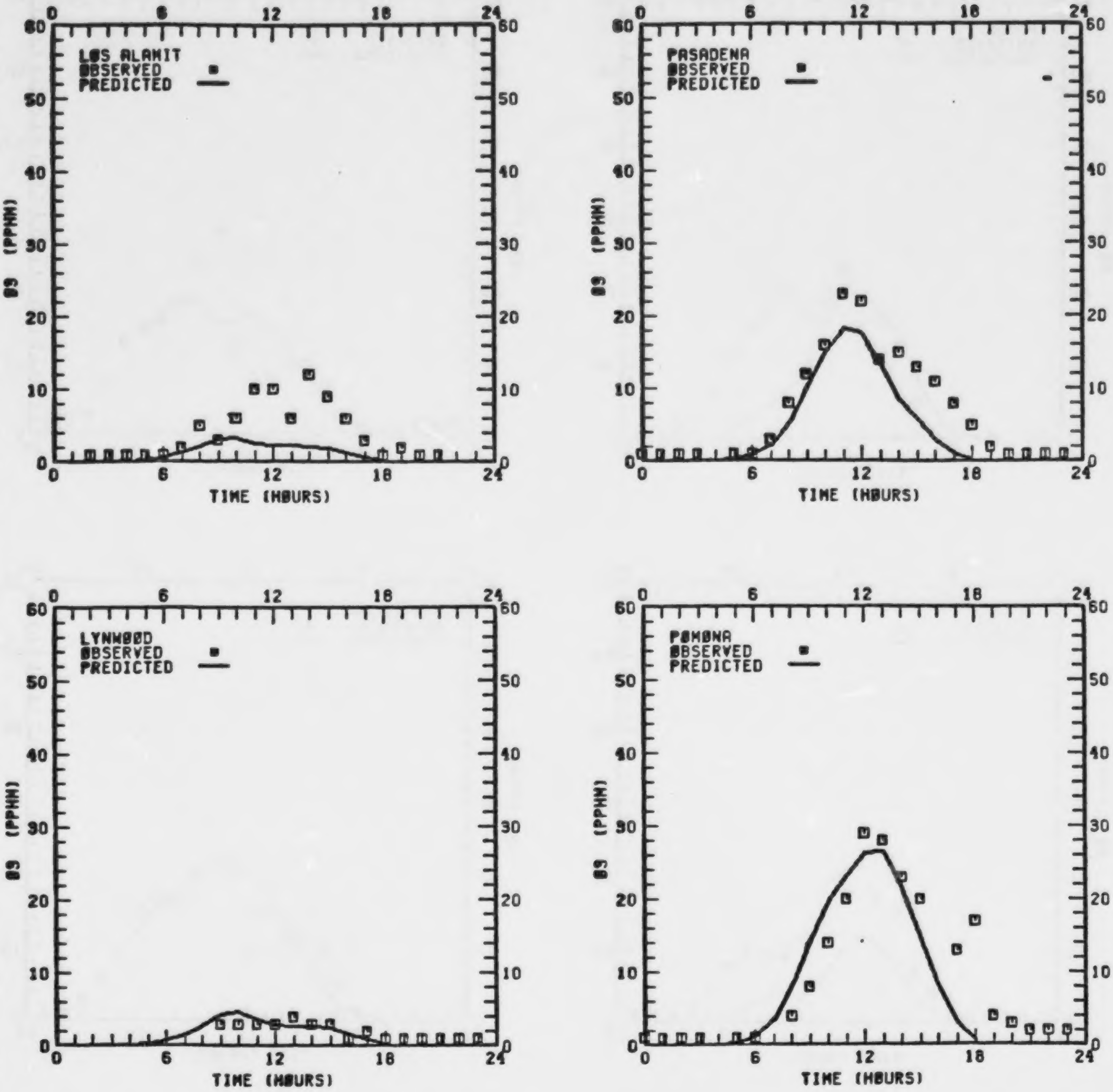


FIGURE IV-7 (Continued)

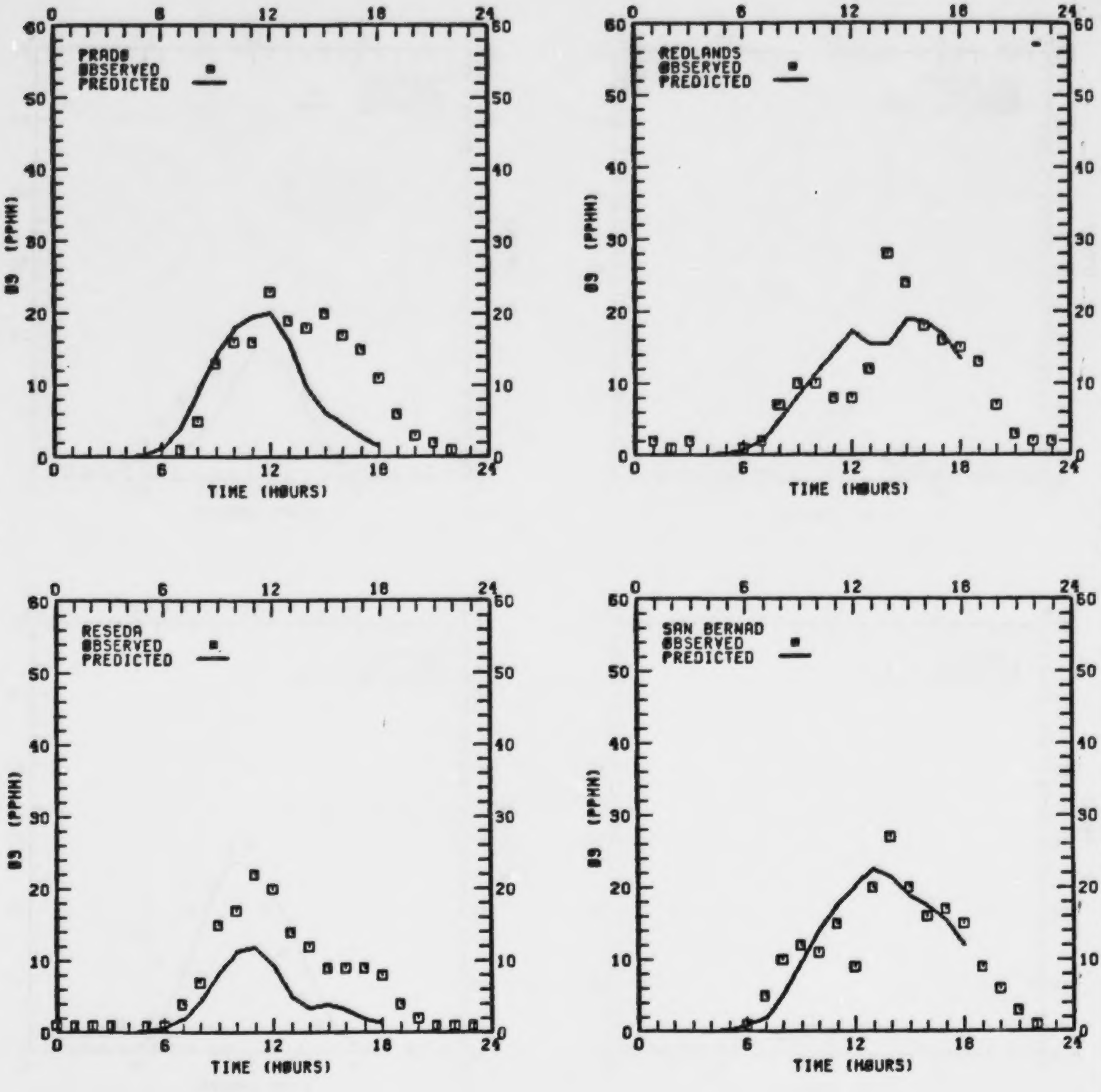


FIGURE IV-7 (Continued)

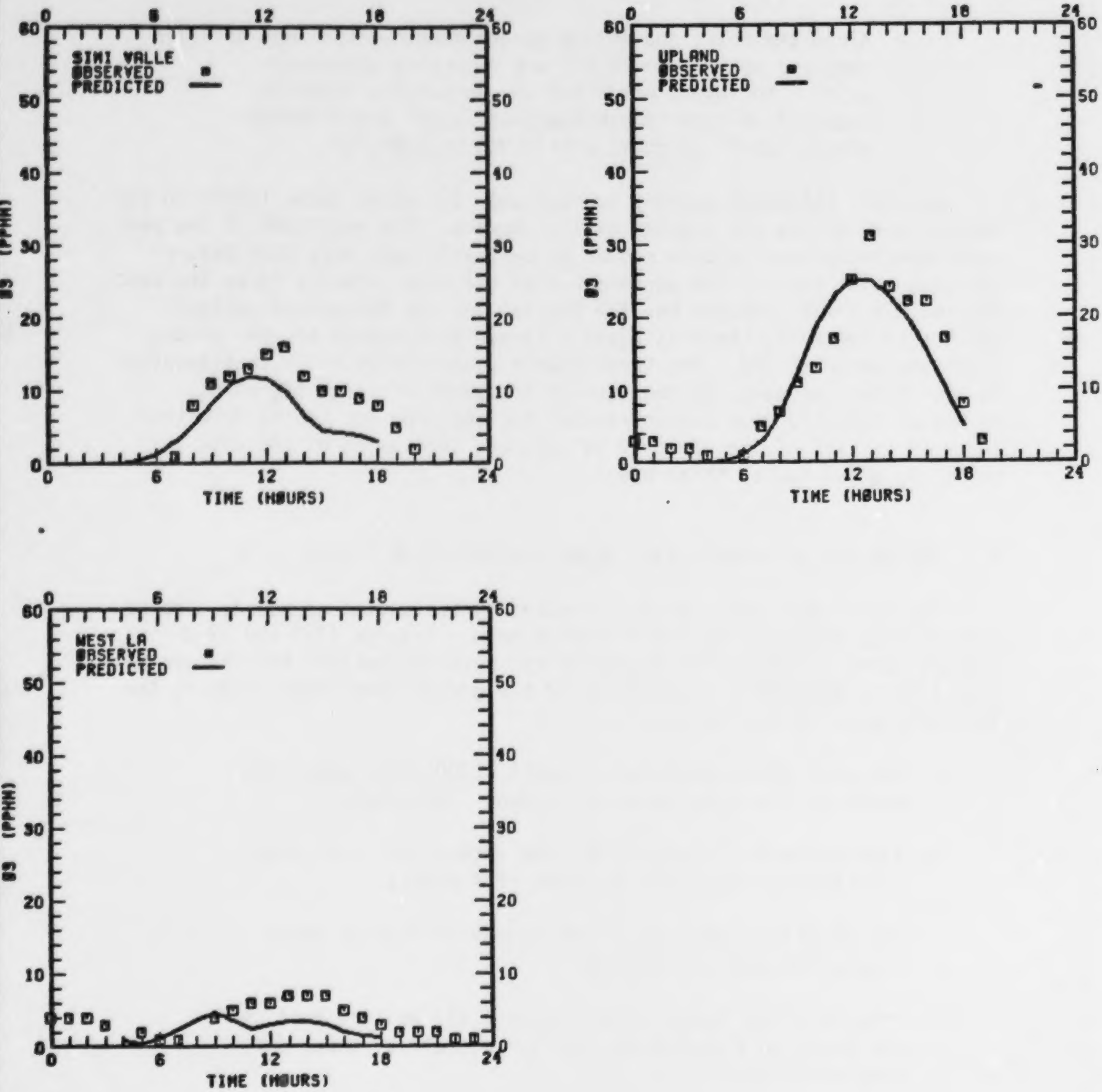


FIGURE IV-7 (Concluded)

- Along the coast and inland toward downtown Los Angeles computed peak ozone levels are typically underestimated. Moreover, predicted concentrations begin to drop before noon though observed values remain moderately high (~ 10 pphm) until 1300 to 1400 PST.

Overall, the model appears to have underestimated ozone levels in the western part of the Los Angeles basin. However, the magnitude of the peak ozone concentrations in this region is typically much less than that of the basin wide peaks. The performance of the model appears to be the best for the mid basin, between the San Gabriel and San Bernardino valleys. Farther to the east, there is again a larger discrepancy between predictions and observations. The trend toward underestimation in that location is attributed, in part, to the current treatment of large NO_x point sources. Computed peak concentrations for the area are typically within 30 to 40 percent of the observed values, and the timing of the peak is offset by about two or three hours.

D. EVALUATION OF GROUND-LEVEL OZONE CONCENTRATION FIELDS

The base case ozone results produced by evaluation of hourly average ground-level ozone fields are presented here. Figures IV-8 and IV-9 present ozone concentration isopleths for both simulations for the period from 1000 to 1600 PST. Examination of the ground-level maps leads to the following general conclusions:

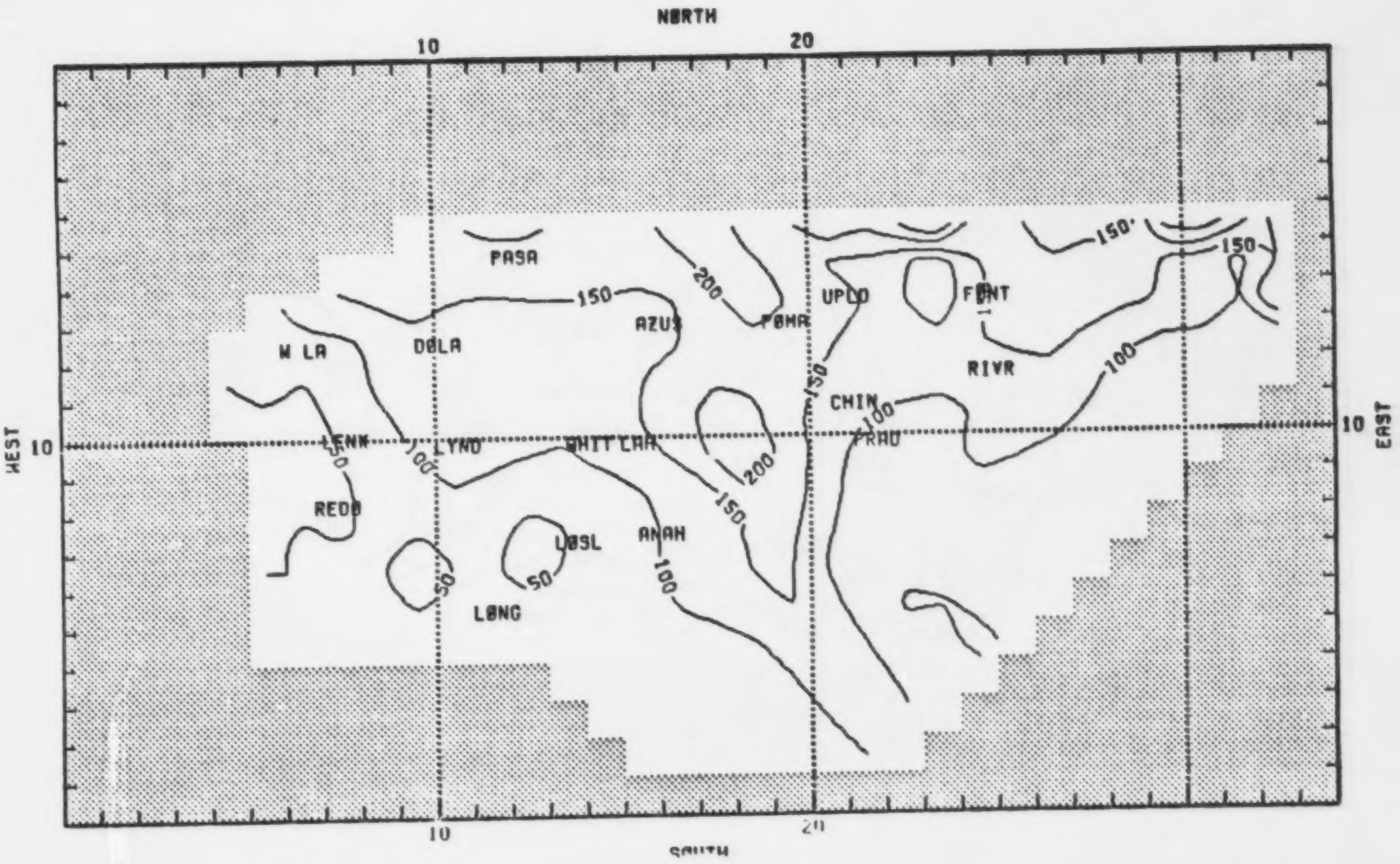
> The peak ozone levels of 4 August (1200-1300) occur two hours earlier than those of 26 June (1400-1500).

> Elevated ozone levels of 26 June extend over a broader geographical area than do those of 4 August.

> The Pomona-Upland area is the region of highest ozone impact for both simulations.

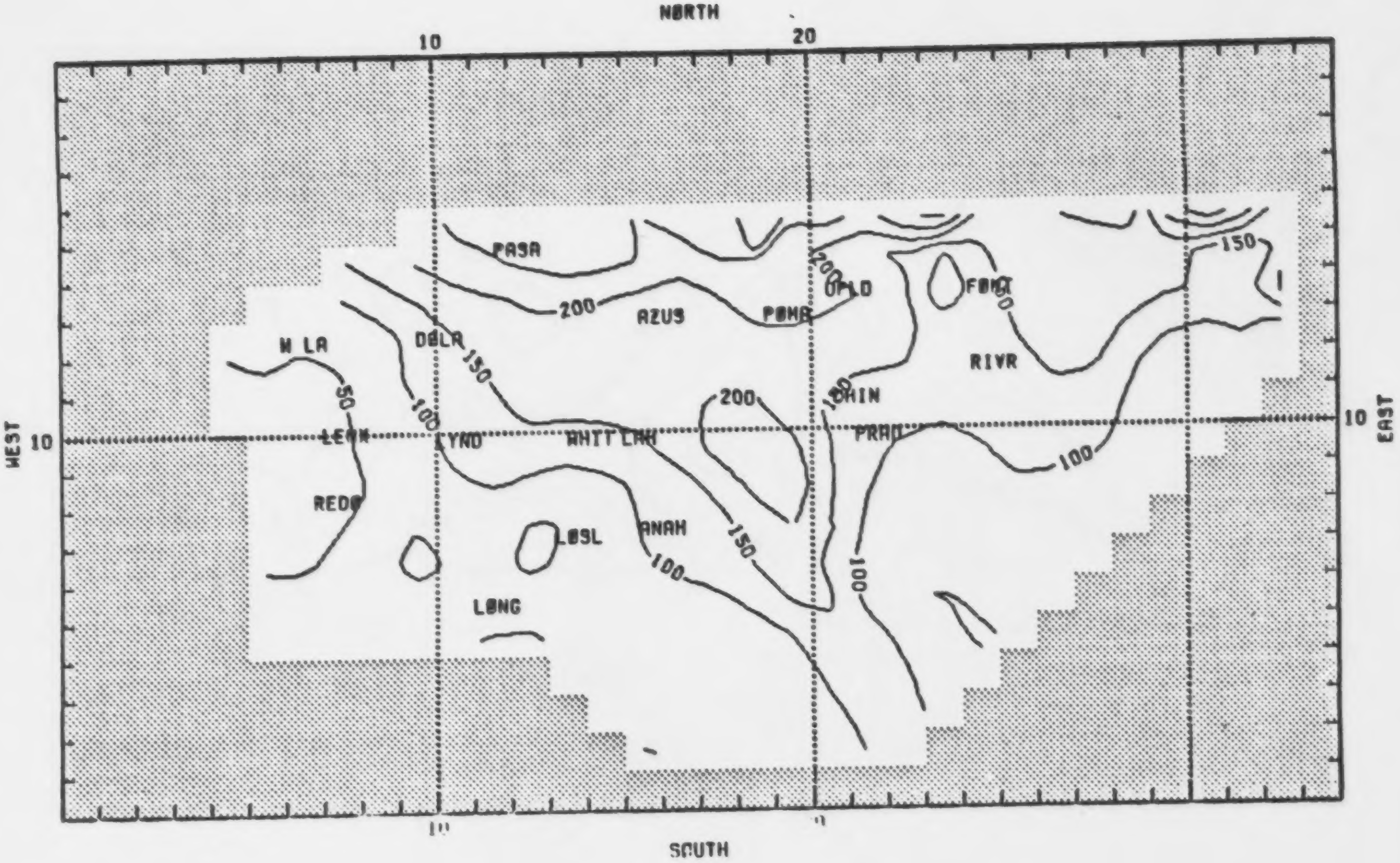
> Concentration levels occurring over the western half of the basin on 4 August are nearly one-half of those occurring on 26 June.

> Both simulations exhibit steep ozone concentration gradients along a SW-NE transect of the San Gabriel and San Bernardino valleys.


NORTH
10
20
WEST
10
EAST
10
SOUTH
10
20
PASA
W LA
DOLA
AZUS
POMA
UPLD
FONT
RIVR
CHIN
PRAD
LENN
LYND
WHIT
LAHA
REDO
LOSL
ANAH
LONG
50
100
150
200

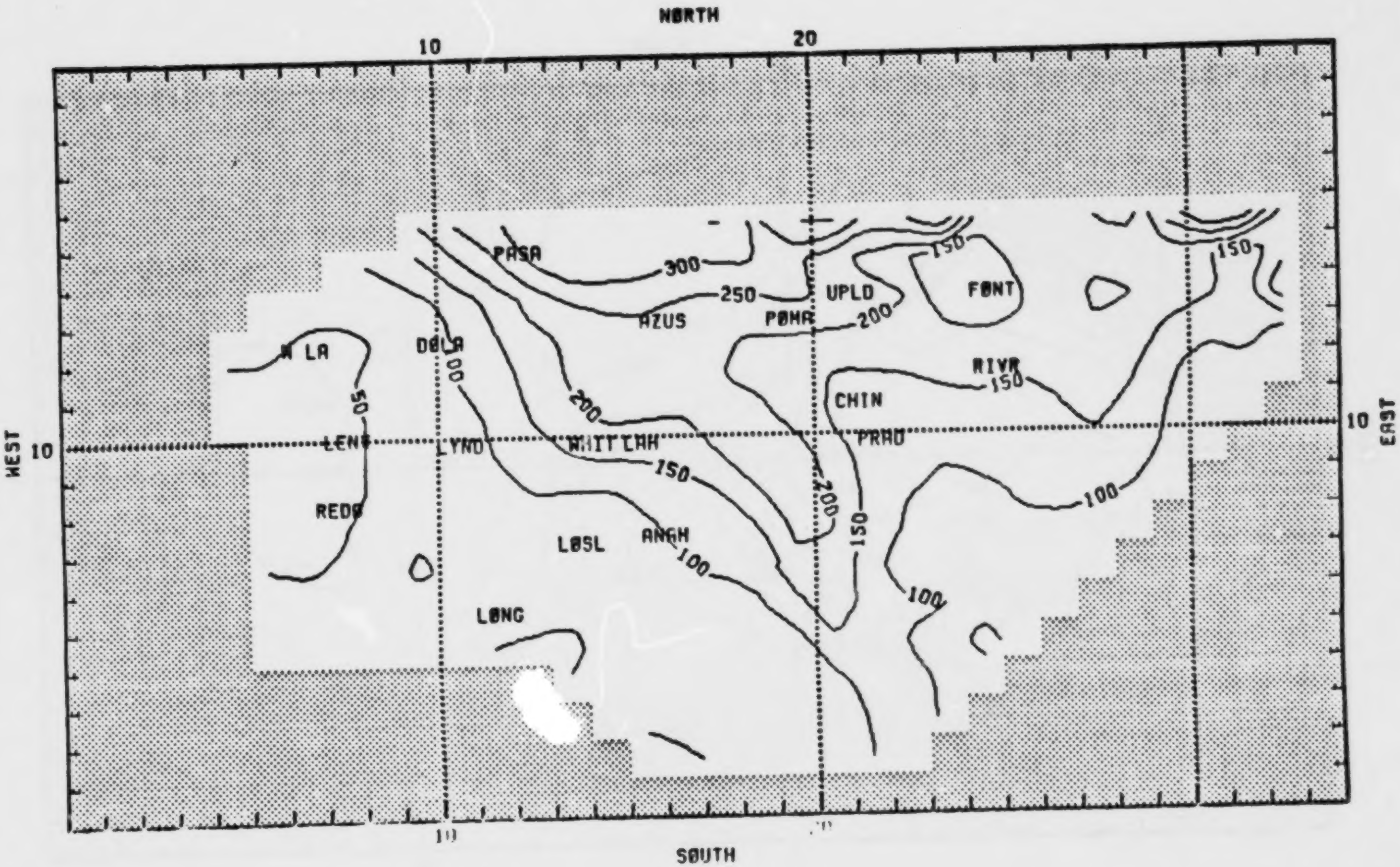
(a) Between the Hours of 10 and 11

FIGURE IV-8. OZONE ISOPLETHS FOR 26 JUNE 1974


NORTH
10
20
WEST
10
EAST
10
SOUTH
PASA
200
AZUS
POMA
200
UPLD
FONT
150
RIVR
W LA
DOLA
150
100
50
LEAN
LYND
WHIT LAH
200
CHIN
PRAD
100
REDO
LBSL
ANAH
100
150
100
LONG

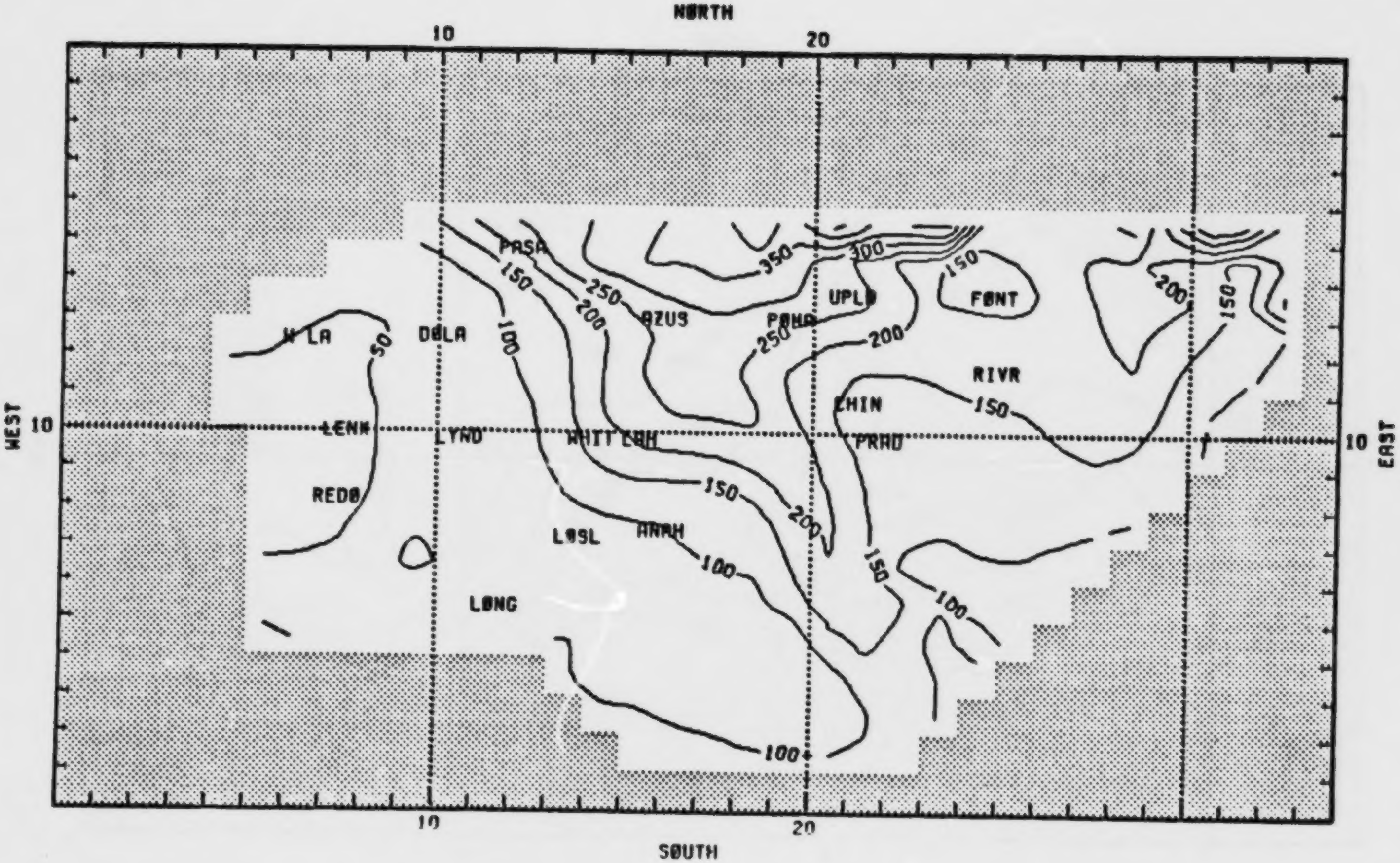
(b) Between the Hours of 11 and 12

FIGURE IV-8 (Continued)


NØRTH
10
20
WEST
10
EAST
10
SØUTH
10
20
PASA
300
250
AZUS
PØMA
UPLD
200
150
FØNT
150
N LA
DØLA
100
50
LENS
LYND
WHIT LAH
200
150
CHIN
PRAD
RIVR
150
100
REDØ
LØSL
ANAH
100
200
150
100
LØNG

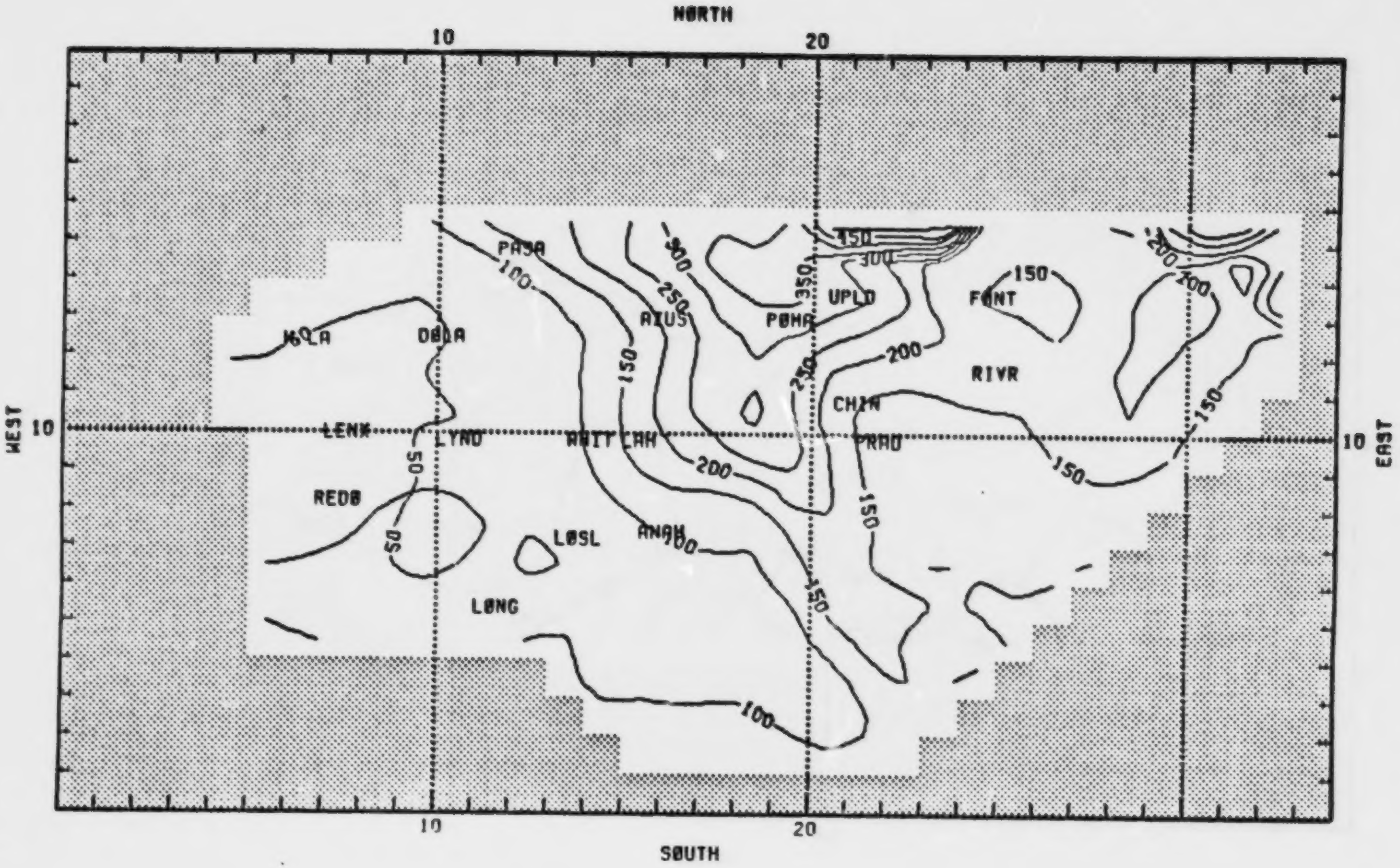
(c) Between the Hours of 12 and 13

FIGURE IV-8 (Continued)


NORTH
10
20
WEST
10
EAST
10
SOUTH
10
20
PASA
AZUS
POMA
UPLD
FONT
W LA
DOLA
RIVR
CHIN
LENN
LYND
WHIT
LAH
PRAD
REDO
LOSL
ANAH
LONG
50
100
150
200
250
350
300
150
200
250
200
150
200
150
150
200
150
100
100
100

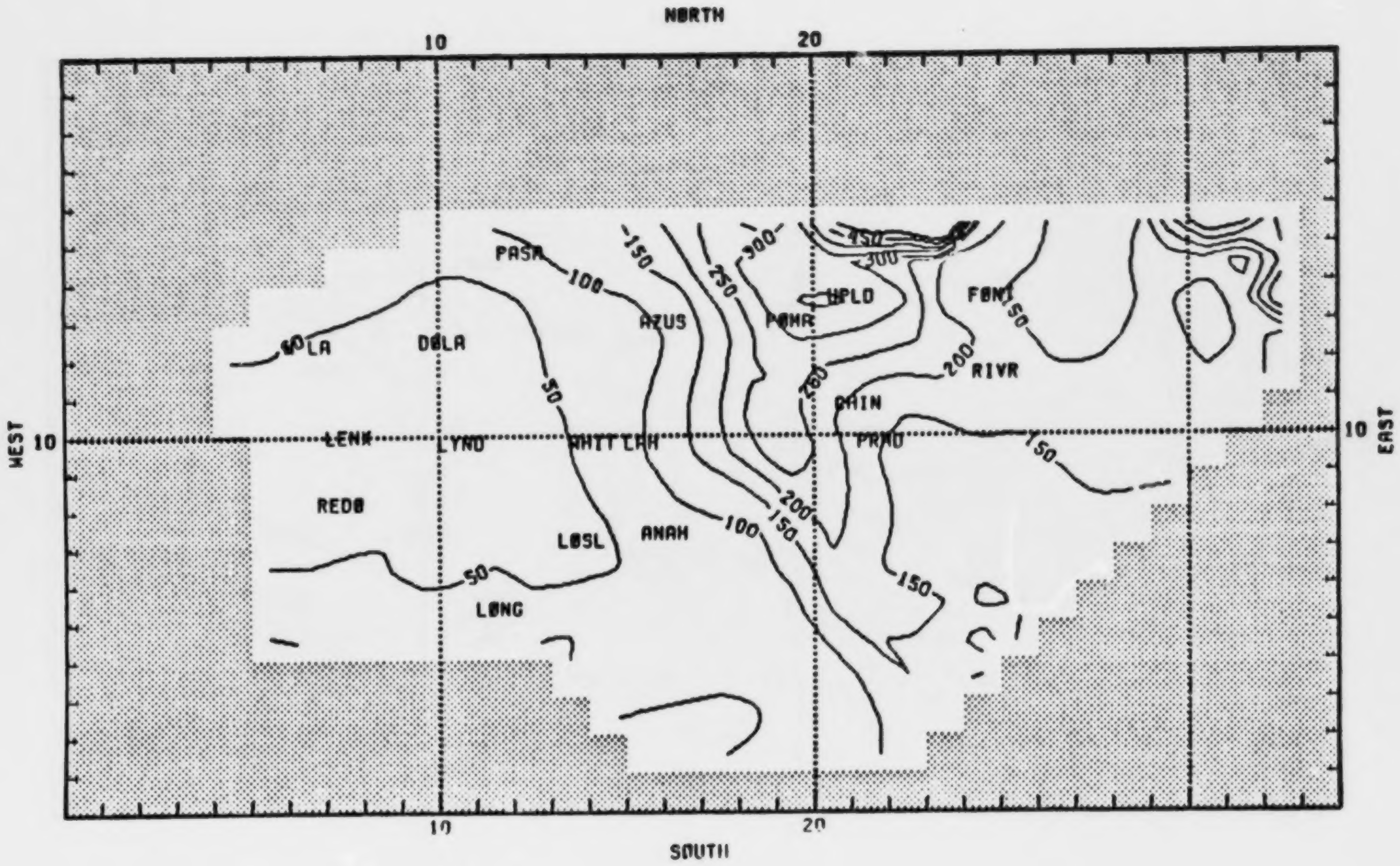
(d) Between the Hours of 13 and 14

FIGURE IV-8 (Continued)


NØRTH
SØUTH
WEST
EAST
10
20
10
20
10
10
PASA
AZUS
PØMA
UPLD
FØNT
RIVR
CHIN
PRAD
WHIT
LAH
ANAH
LØSL
LØNG
REDB
LYND
LENX
DØLA
WSLA
50
100
150
200
250
300
350

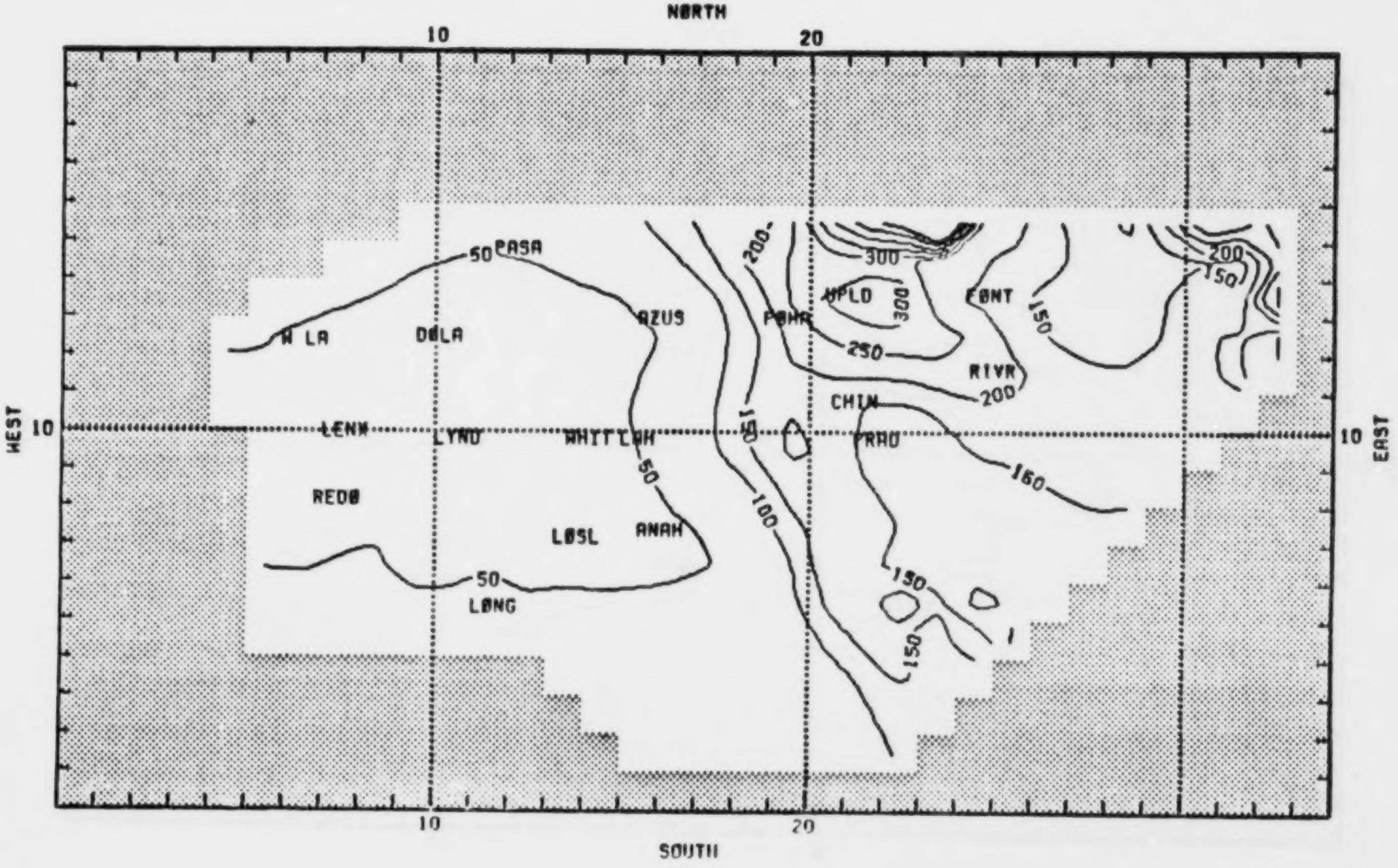
(e) Between the Hours of 14 and 15

FIGURE IV-8 (Continued)


NORTH
10
20
WEST
10
EAST
10
SOUTH
10
20
PASA
100
-150
250
300
450
300
UPLD
FONT
150
LA
50
DOLA
AZUS
POMR
200
RIVR
50
250
CHIN
LENX
LYND
WHITT
LAH
PRAD
150
REDB
200
LOSL
ANAH
100
150
50
150
LONG

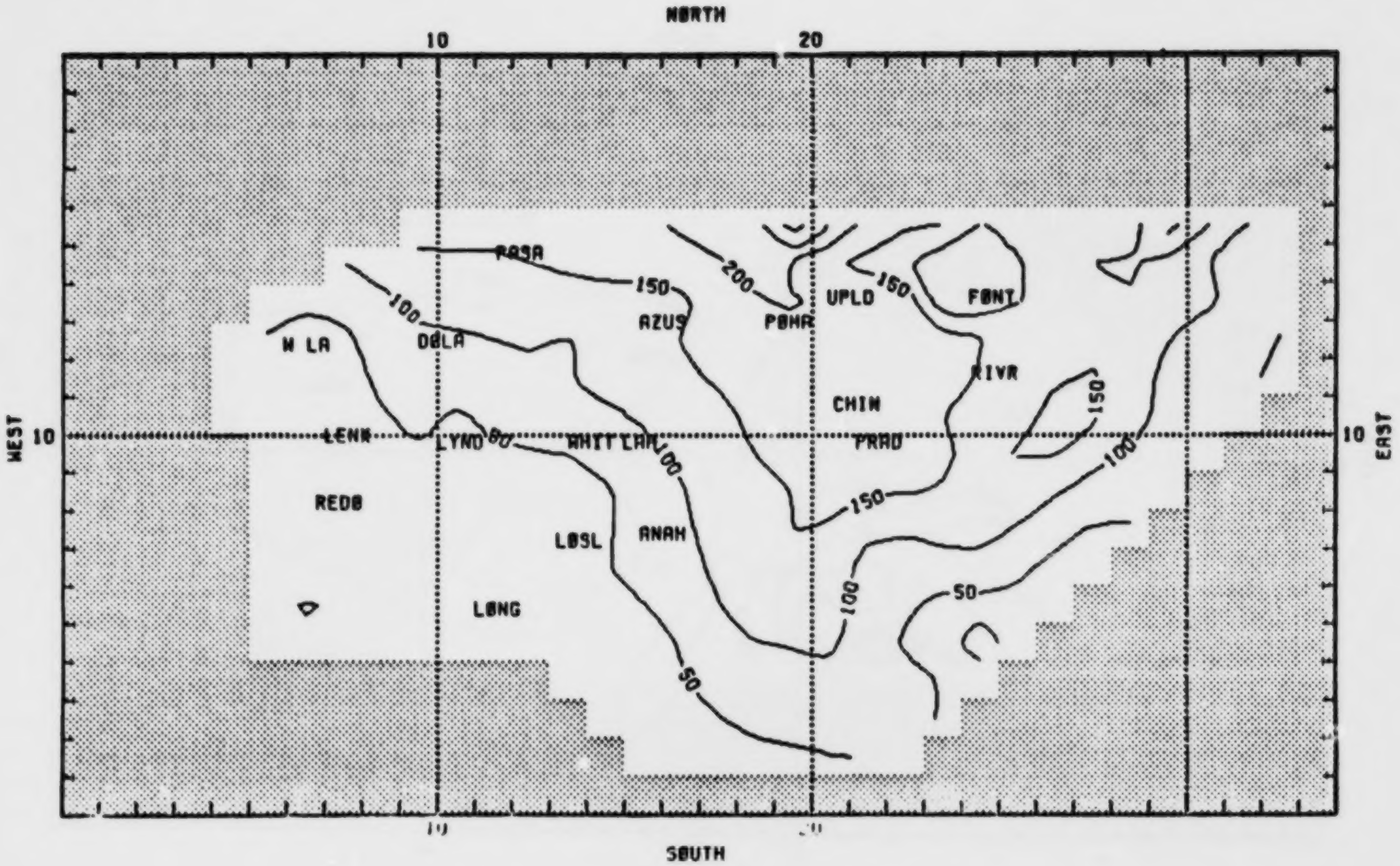
(f) Between the Hours of 15 and 16

FIGURE IV-8 (Continued)


NORTH
SOUTH
WEST
EAST
10
20
10
20
10
10
W LA
DOLA
PASA
AZUS
POMA
UPLD
FONT
RIVR
CHIN
PRAD
LENX
LYND
WHIT
LAH
REDO
LOSL
ANAH
LONG
50
50
50
100
150
150
150
150
150
150
160
200
200
200
250
300
300

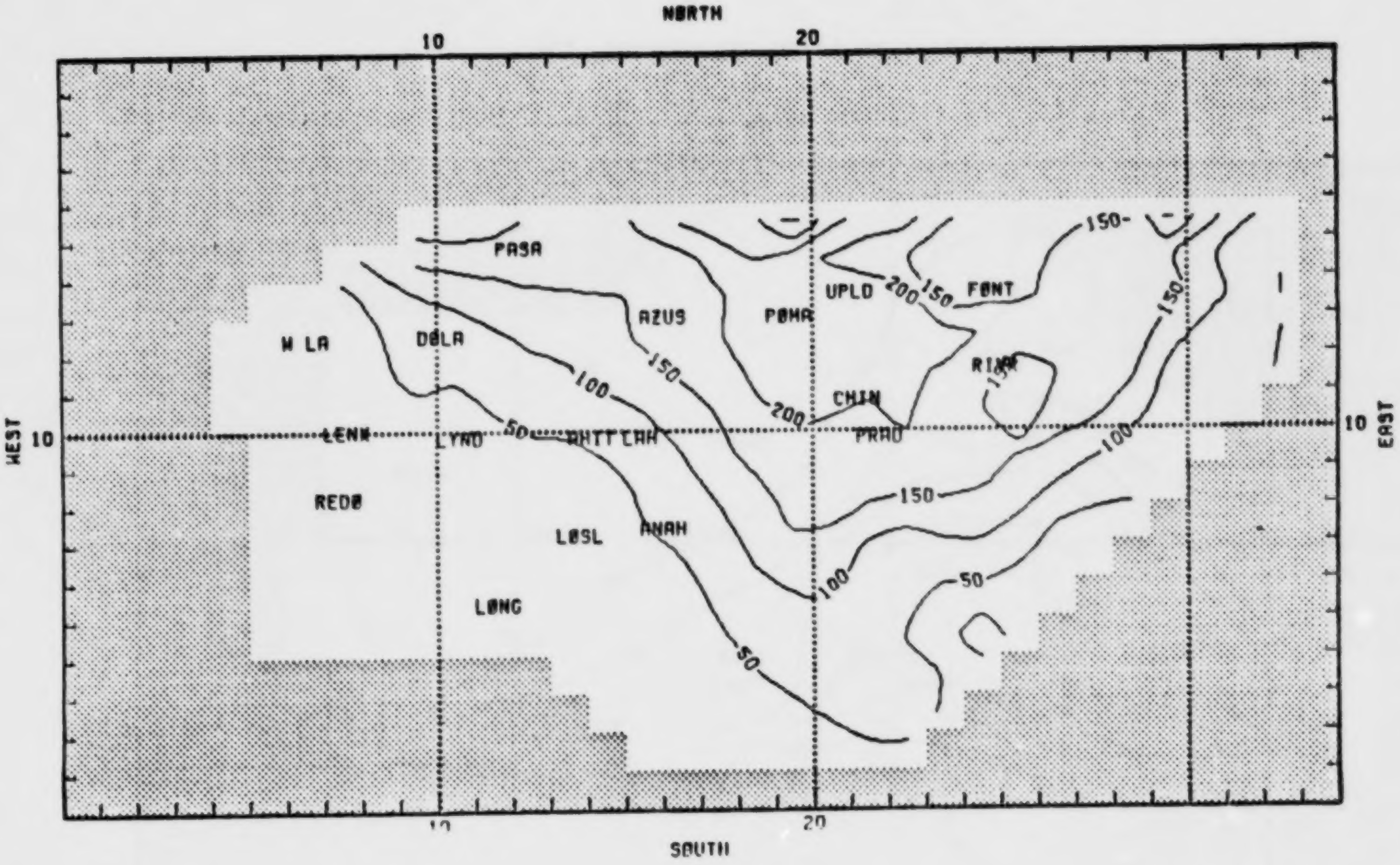
(g) Between the Hours of 16 and 17

FIGURE IV-8 (Concluded)


NORTH
SOUTH
WEST
EAST
10
20
10
10
W LA
DOLA
PASA
AZUS
POMA
UPLD
FONT
RIVR
CHIN
PRAD
LENN
LYND
WHITLAH
REDB
LOSL
ANAH
LONG
50
100
150
200

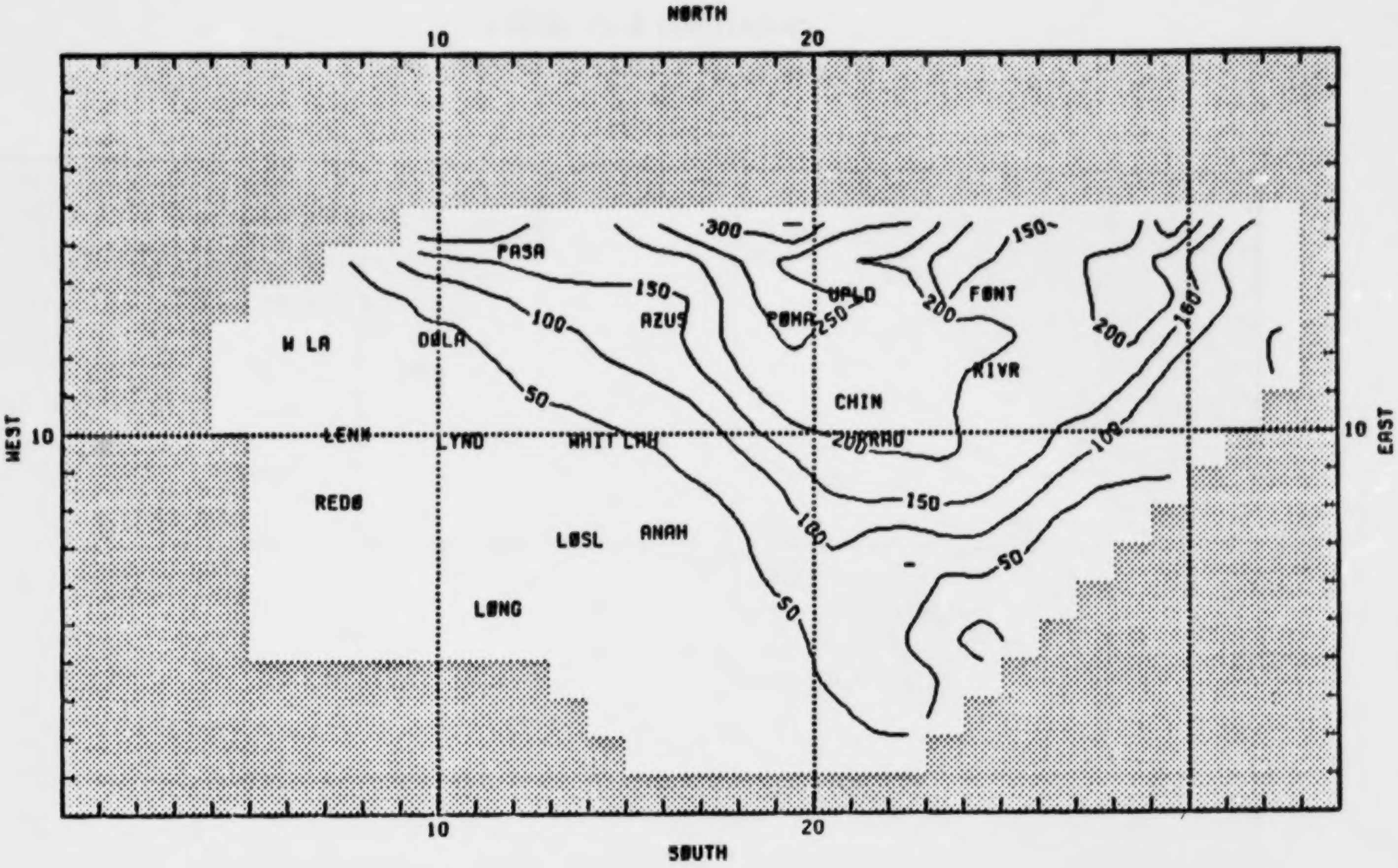
(a) Between the Hours of 10 and 11

FIGURE IV-9. OZONE ISOPLETHS FOR 4 AUGUST 1975


NORTH
SOUTH
WEST
EAST
10
20
10
20
10
10
PASA
AZUS
POMA
UPLD
FONT
W LA
DOLA
RIVR
CHIN
LENX
LYND
WHTT
LAH
PRAD
REDO
LOSL
ANAH
LONG
50
100
150
200
150
150
200
150
150
150
100
150
100
50
50
100

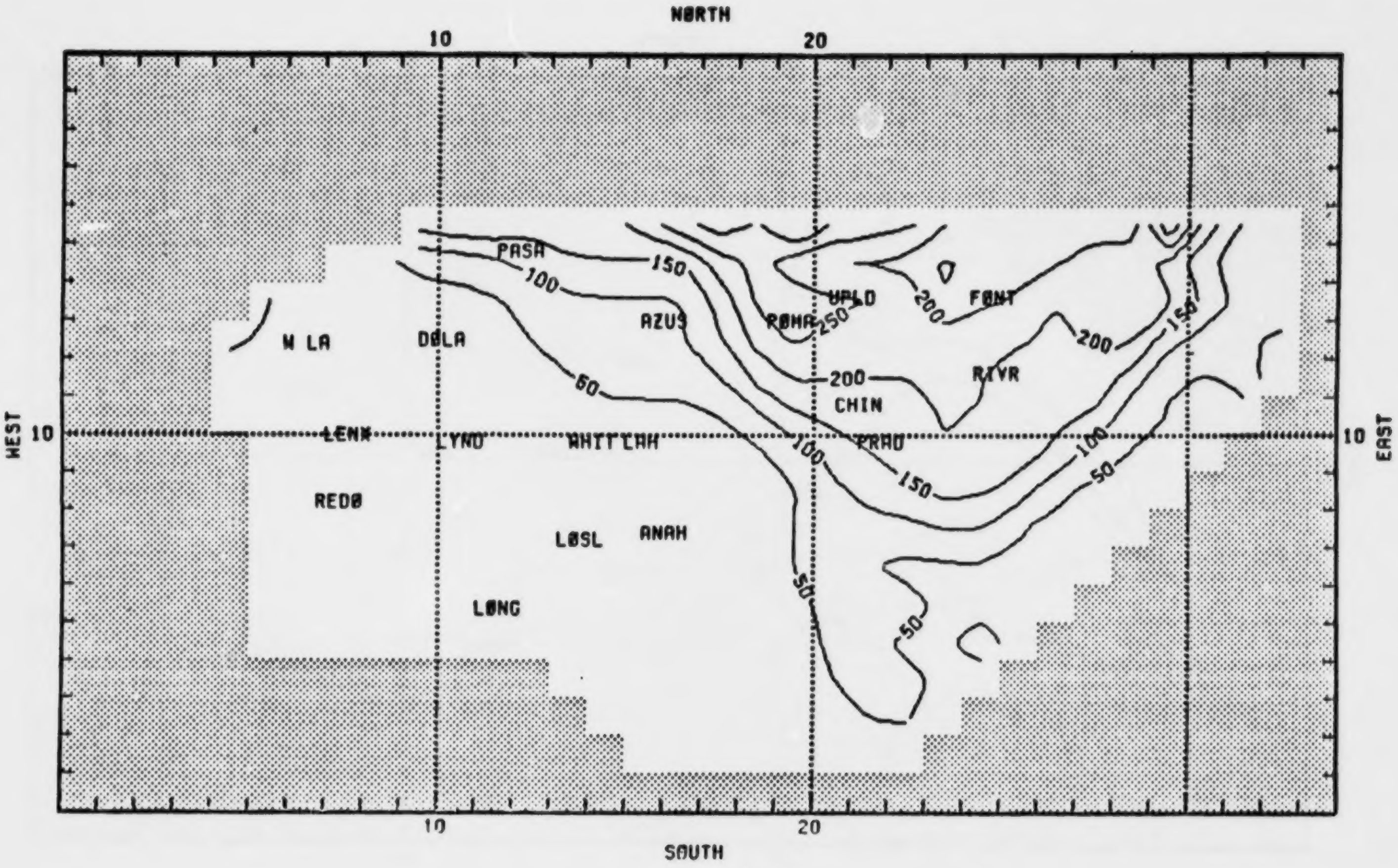
(b) Between the Hours of 11 and 12

FIGURE IV-9 (Continued)


NORTH
SOUTH
WEST
EAST
10
20
10
20
10
10
PASA
W LA
DOLA
AZUS
POMA
UPLD
FONT
RIVR
CHIN
LENN
LYND
WHIT
LAH
PRAD
REDO
LOSL
ANAH
LONG
300
250
200
200
200
160
150
150
150
100
100
100
50
50
50

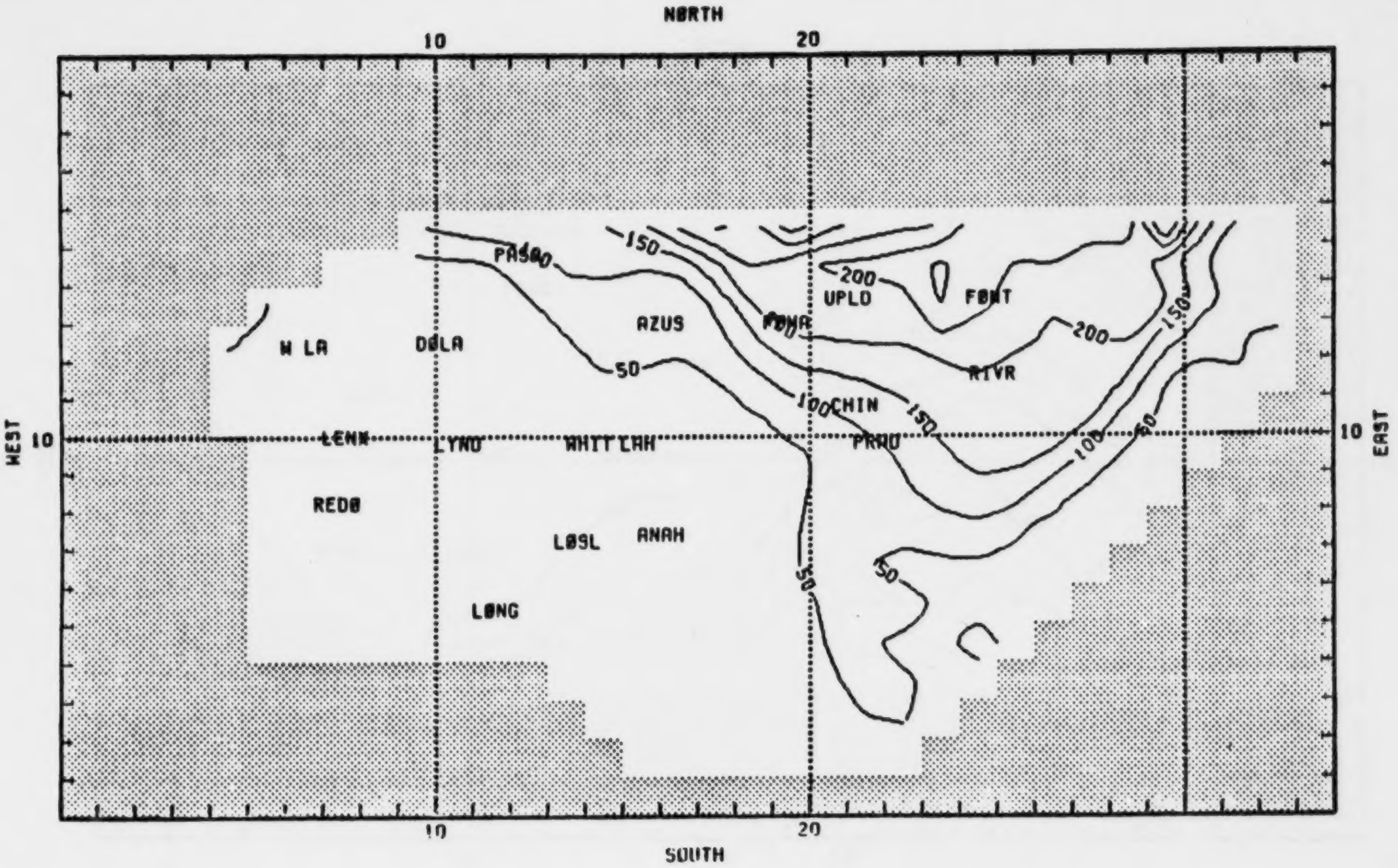
(c) Between the Hours of 12 and 13

FIGURE IV-9 (Continued)


NORTH
SOUTH
WEST
EAST
10
20
10
20
10
10
W LA
DOLA
PASA
AZUS
POMA
UPLD
FONT
RIVR
CHIN
PRAD
LENX
LYND
WHITTLAH
REDO
LOSL
ANAH
LONG
50
100
150
200
250

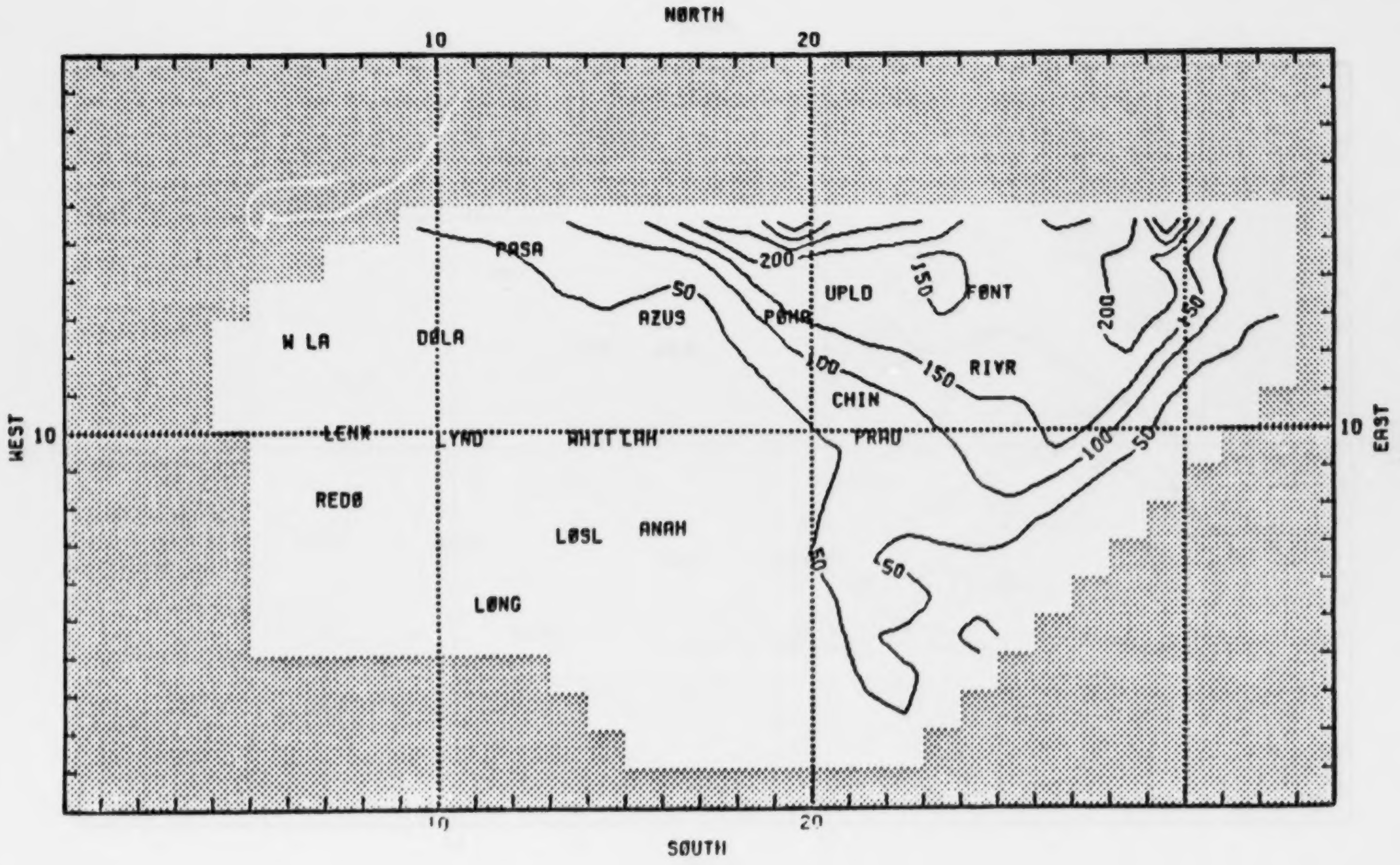
(d) Between the Hours of 13 and 14

FIGURE IV-9 (Continued)


NORTH
SOUTH
WEST
EAST
10
20
10
20
10
10
W LA
DOLA
PASAD
AZUS
POMA
UPLD
FONT
RIVR
CHIN
PRAD
LENX
LYND
WHIT
LAH
REDO
LOSL
ANAH
LONG
50
100
150
200
200
150
100
150
100
50
50
50

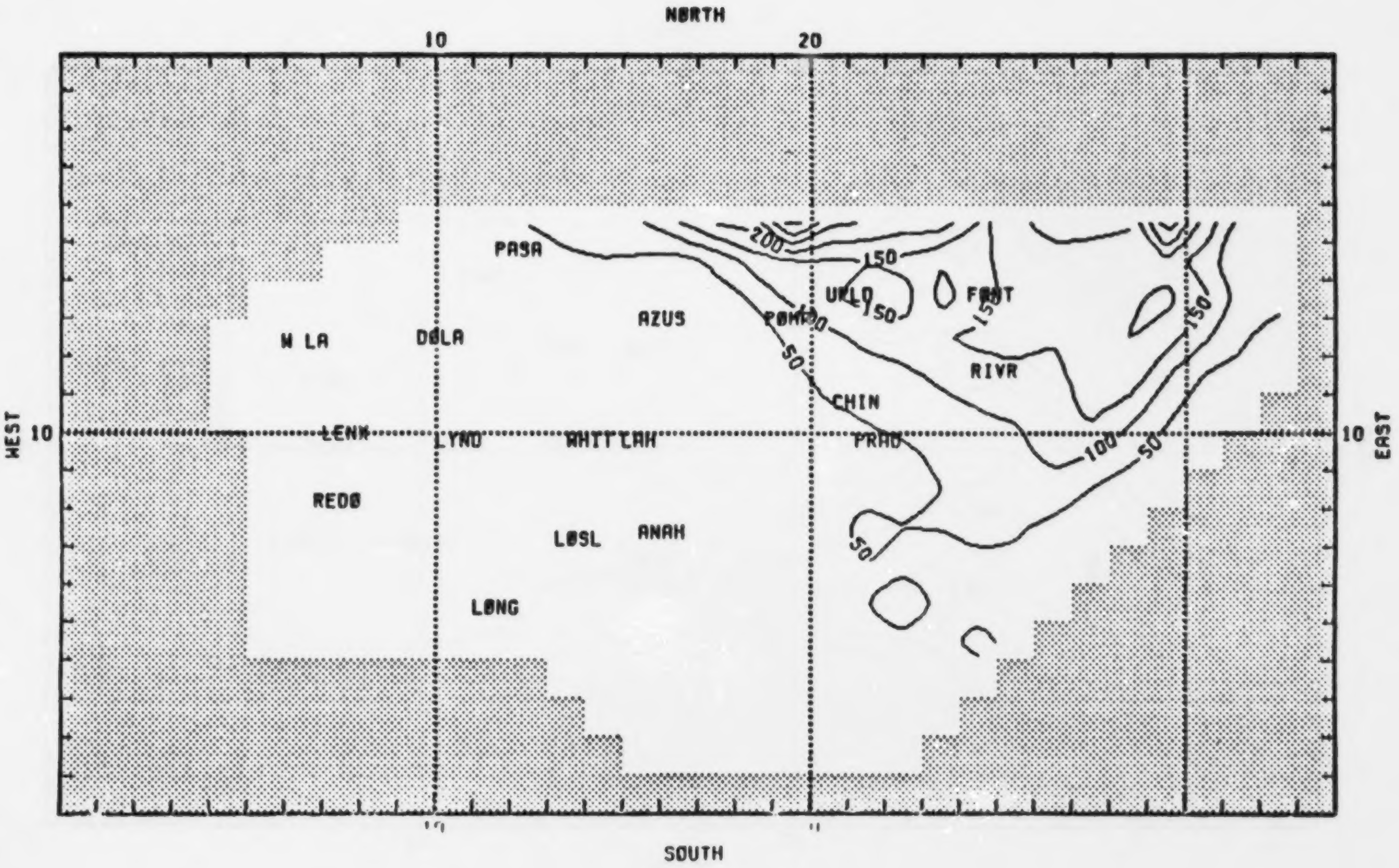
(e) Between the Hours of 14 and 15

FIGURE IV-9 (Continued)


NORTH
10
20
WEST
10
EAST
10
SOUTH
10
20
PASA
AZUS
POMA
UPLD
FONT
RIVR
CHIN
PRAD
W LA
DOLA
LENX
LYND
WHIT LAH
REDO
LOSL
ANAH
LONG
50
100
150
200
150
200
150
100
50
50
50

(f) Between the Hours of 15 and 16

FIGURE IV-9 (Continued


NØRTH
10
20
WEST
10
EAST
10
SOUTH
PASA
W LA
DØLA
AZUS
PØMA
UPLD
FØNT
RIVR
CHIN
PRAD
LENX
LYND
WHITTLAH
REDØ
LØSL
ANAH
LØNG
200
150
150
150
150
100
100
50
50
50
50

(g) Between the Hours of 16 and 17

FIGURE IV-9 (Concluded)

E. SUMMARY

This chapter presents the results of the 26 June 1974 and 4 August 1975 base case airshed model simulations. Overall, both sets of results agree favorably with ozone observations.

V SIMULATION RESULTS

A. INTRODUCTION

Figure V-1 presents schematically the relationship between input data, model input variables, and air quality model predictions. The objective of this study is to determine the sensitivity of airshed model ozone predictions to the amount (or "richness") of input data. This objective clearly differs from that of classical sensitivity analysis studies. The latter deal with the sensitivity of a mathematical model's predictions to variations in model parameters. To place this study in perspective, various approaches to sensitivity analyses of air quality models are summarized. The discussion then shifts to how the sensitivity of the airshed model to various levels of detail in input data is ascertained.

1. Review of Sensitivity Analyses of Air Quality Models

Mathematical model sensitivity analysis may be defined as the evaluation of the deviation in model output resulting from a perturbation in one or several model inputs. One possible approach consists in defining a mathematical relationship that relates the extent of the model output deviation to the perturbation in the input variables (or parameters). Such sensitivity analysis methods have been applied to some air pollution models (Seigneur, 1978; Koda et al., 1979; McRae and Tilden, 1980); however, their application to complex urban air quality models would be too costly and cannot be considered at present.

Another approach that has been widely used consists in perturbing the model input variables and parameters one by one. This approach gives less information (on a "per run" basis) than the more complex mathematical methods mentioned above, but the information obtained, if used appropriately, may be very useful for understanding the dynamics of the model. For example, Liu et al. (1976) and Dunker (1980) used this approach to evaluate the sensitivity of the airshed model predictions to perturbation in emissions, wind velocities, vertical eddy diffusivity and mixing height. The same technique was also used to study the sensitivity of atmospheric chemical mechanisms to individual reaction rates (Dodge and Hecht, 1975; Duewer et al., 1977).

In this study, the latter approach was necessary because of the complexity of the urban air quality model. Specific perturbations have been introduced in the input data, and the resulting changes in the model predictions provide some information on the model sensitivity.

2. Sensitivity of the Airshed Model to Input Data

A detailed description of the airshed model is reported by Reynolds et al. (1973). A brief overview can be described as follows.

The airshed model consists of a simulation model and several submodels that supply various input variables and parameters. Figure V-1 shows that the simulation model is comprised of a set of continuity equations that describe the advective transport, turbulent dispersion, chemical transformations, pollutant emissions, and removal of chemical species in the atmosphere. Model input variables such as wind fields, eddy diffusivities, reaction-rate coefficients, and emission rates are defined either directly or by means of various submodels. Initial and boundary conditions are estimated from air quality data and historical intensive field measurement programs. The airshed model configuration can vary in terms of grid size and number of grid layers. Although there are basic theoretical constraints on the grid size (Lamb and Seinfeld, 1973; Reynolds et al., 1973), these features may be varied according to the airshed considered.

As suggested in figure V-1, analysis of airshed model sensitivity to input data may be considered in two parts: sensitivity of the model output to changes in the input variables (e.g., wind fields, diffusivities, emission rates, mixing depths) may be evaluated through successive model simulations, each involving a prescribed set of input files; and the sensitivity of these input variables to perturbations in the input data used to construct them. For instance, wind field maps will show the effect of a change in meteorological data on wind speeds and direction. Attention is therefore focused on (1) the combination of the sensitivity of the model to its input variables and on (2) the sensitivity of the input variables to the basic data.

B. MEASURES FOR ASCERTAINING MODEL SENSITIVITY

An important step in quantifying model sensitivity is the definition of specific measures. Several measures that appear to be useful include the following:

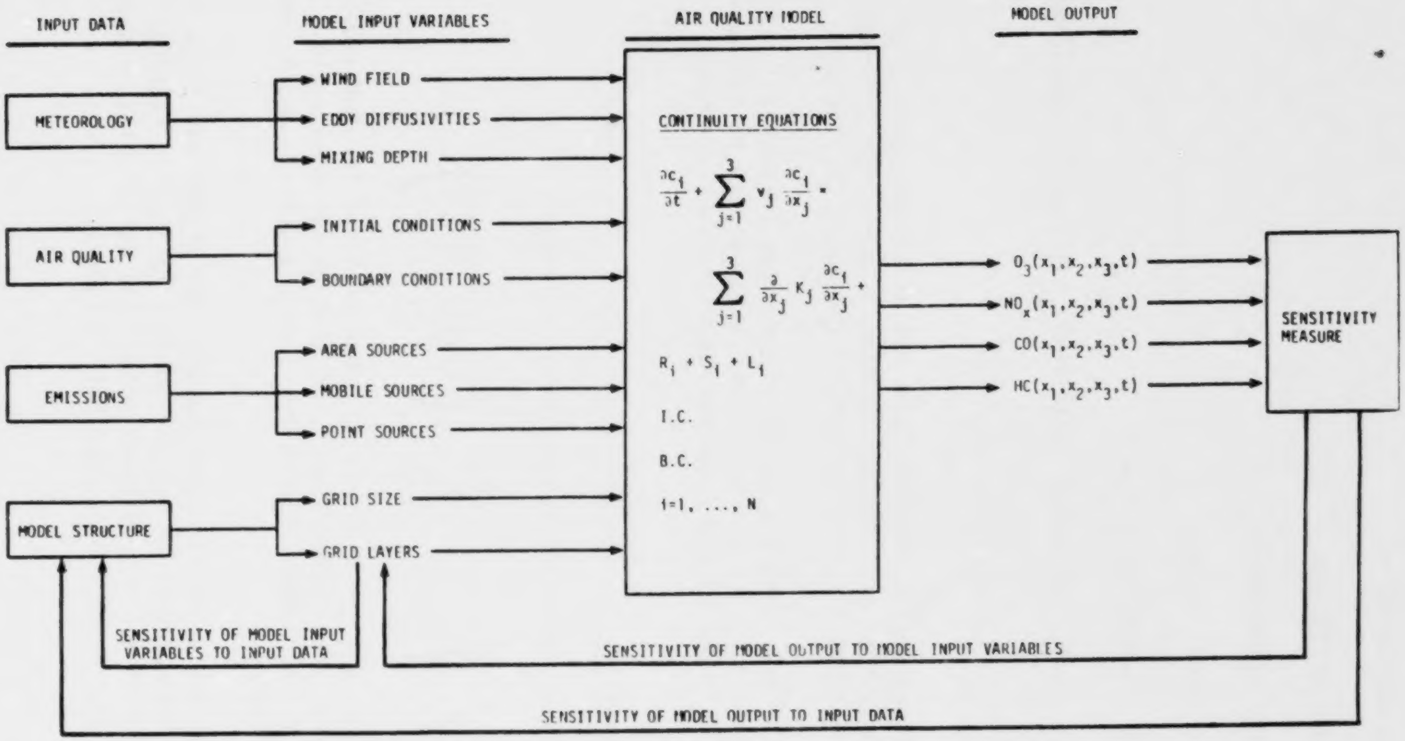

INPUT DATA
MODEL INPUT VARIABLES
AIR QUALITY MODEL
MODEL OUTPUT
METEOROLOGY
AIR QUALITY
EMISSIONS
MODEL STRUCTURE
WIND FIELD
EDDY DIFFUSIVITIES
MIXING DEPTH
INITIAL CONDITIONS
BOUNDARY CONDITIONS
AREA SOURCES
MOBILE SOURCES
POINT SOURCES
GRID SIZE
GRID LAYERS
CONTINUITY EQUATIONS
$\frac{\partial c_i}{\partial t} + \sum_{j=1}^{3} v_j \frac{\partial c_i}{\partial x_j} = \sum_{j=1}^{3} \frac{\partial}{\partial x_j} K_j \frac{\partial c_i}{\partial x_j} + R_i + S_i + L_i$
I.C.
B.C.
$i=1, \ldots, N$
$O_3(x_1,x_2,x_3,t)$
$NO_x(x_1,x_2,x_3,t)$
$CO(x_1,x_2,x_3,t)$
$HC(x_1,x_2,x_3,t)$
SENSITIVITY MEASURE
SENSITIVITY OF MODEL INPUT VARIABLES TO INPUT DATA
SENSITIVITY OF MODEL OUTPUT TO MODEL INPUT VARIABLES
SENSITIVITY OF MODEL OUTPUT TO INPUT DATA

FIGURE V-1. SENSITIVITY ANALYSIS OF THE URBAN AIRSHED MODEL

> Signed deviation.

> Absolute deviation.

> Temporal correlation.

> Spatial correlation.

> Overall maximum ozone level.

> Maximum ozone statistics (peak level difference, peak time lag).

> Dosage.

In addition, isopleths of maximum ozone deviation and ozone profiles at air quality monitoring stations are developed to provide information on the spatial and temporal perturbations in the model predictions.

Note that these measures, to be defined more precisely in the following subsections, are quite similar to those used to evaluate model performance in simulating the June and August oxidant episodes.

1. Signed Deviation

The signed deviation is calculated as follows:

$$\frac{1}{N'} \sum_{j=1}^{N'} \frac{1}{N} \sum_{i=1}^{N} \left(\frac{C_{s,i,j} - C_{b,i,j}}{C_{b,i,j}} \right) , \qquad (5\text{-}1)$$

where $C_{s,i,j}$ and $C_{b,i,j}$ are the ozone concentrations for the sensitivity case and the base case, respectively, at station (or grid cell) i and for the hour j; N is the number of stations (or grid cells), and N' the number of simulation hours.

Two signed deviations are calculated. The signed deviation is computed for the ozone levels at the air quality monitoring stations and for the ozone levels in all ground-level grid cells (grid cells in mountainous areas and over the Pacific ocean are not included).

2. Absolute Deviation

The absolute deviation in ozone levels is computed as follows:

$$\frac{1}{N'} \sum_{j=1}^{N'} \frac{1}{N} \sum_{i=1}^{N} \left| \frac{C_{s,i,j} - C_{b,i,j}}{C_{b,i,j}} \right| \quad . \tag{5-2}$$

Absolute deviation in ozone concentration levels (i.e., base case vs. simulation) is computed for the air quality monitoring stations and for the ground-level grid cells.

3. Temporal Correlation

The temporal correlation refers to the "timing" of the ozone concentration levels computed by both the sensitivity case and the base case at a specified station or in a specified grid cell. The temporal correlation at a given location is determined from the hourly concentrations predicted by the sensitivity case and the base case at a given location j. A correlation coefficient is then calculated for each station according to routine statistics. These correlation coefficients are normalized with respect to the "perfect correlation line" (Hoel, 1962) by performing the following change of variable:

$$\phi_j = \frac{1}{2} \ln \frac{1 + r_j}{1 - r_j} \quad , \tag{5-3}$$

where r_j is the computed correlation coefficient for the station or grid cell j. The mean value of the ϕ_j's is computed for all locations:

$$\overline{\Phi} = \frac{1}{N} \sum_{j=1}^{N} \phi_j \quad , \tag{5-4}$$

where N is the number of stations or grid cells. Because the values of ϕ_j are approximately normally distributed (Hayes, 1978), the average temporal correlation coefficient ρ is evaluated from the following formula:

$$\overline{\Phi} = \frac{1}{2} \ln \frac{1 + \rho}{1 - \rho} \quad . \tag{5-5}$$

Thus, the average temporal correlation coefficient is:

$$\rho = \frac{\exp\,(2\,\overline{\Phi}) - 1}{\exp\,(2\,\overline{\Phi}) + 1} \quad . \qquad (5\text{-}6)$$

The temporal correlation is computed both from station and grid statistics. Perfect correlation exists when $\rho = 1.0$.

4. Spatial Correlation

The spatial correlation between the concentration fields calculated in a sensitivity run and those calculated in the base case is another useful measure. Hourly correlation coefficients can be computed by considering the values of concentrations predicted by the sensitivity case and by the corresponding base case for each station or grid cell. Then, the estimation of the average spatial correlation coefficient follows the procedure described above for the temporal correlation coefficient. Two spatial correlation coefficients are computed from station statistics and from grid statistics, respectively.

These sensitivity measures (signed deviation, absolute deviation, temporal correlation, and spatial correlation) may be evaluated as a function of concentration level. They are computed in this study for ozone levels > 12 pphm (the National Ambient Air Quality Standard) and for ozone levels > 20 pphm.

5. Overall Maximum Ozone Level

The maximum ground level ozone concentration in the basin is computed for the sensitivity case and the base case from station and grid statistics. The overall maximum ozone level and the corresponding location (station or grid) provide an additional measure of model sensitivity.

6. Maximum Ozone Statistics

The maximum ozone levels occurring at each monitoring station may be computed for both the sensitivity and the base case. The average peak level normalized ozone difference may be defined as follows:

$$\frac{1}{N} \sum_{i=1}^{N} \left| \frac{C^{*}_{s,i} - C^{*}_{b,i}}{C^{*}_{b,i}} \right| \quad , \qquad (5\text{-}7)$$

where $C^*_{s,i}$ and $C^*_{b,i}$ are the maximum ozone concentrations for the sensitivity case and the base case, respectively, at Station i. This measure is computed for the 23 air quality monitoring stations within the computational grid. Coastal stations are principally affected by boundary conditions and local emissions since the wind flow in both base cases is westerly for the majority of the simulation time. Accordingly, these stations are removed from calculation of the maximum ozone statistics.* Peak level normalized differences are computed for the remaining 16 downwind stations.

The average absolute peak time lag for maximum ozone level occurrence is evaluated as follows:

$$\frac{1}{N} \sum_{i=1}^{N} \left| T_s - T_b \right| \quad , \tag{5-8}$$

where T_s and T_b are the time of the maximum ozone level in the sensitivity case and in the base case, respectively. This measure is evaluated for the 23 stations and for the 16 downwind stations.

7. Dosage

Dosage measures were based on the gridded area with simulated ozone levels above 20 pphm. Dosage is obtained by adding the number of grid cells with ozone levels above 20 pphm over the entire 19-hour simulation period. This measure is computed for both the base cases and sensitivity cases. The normalized difference of dosages provides an additional measure of the model sensitivity:

$$\text{Normalized difference} = \left(\frac{\text{Dosage of sensitivity case - Dosage of base case}}{\text{Dosage of base case}} \right) \quad .$$

8. Isopleths of Maximum Ozone Deviation

Quantification of the spatial changes in predicted ozone levels aids in the interpretation of the sensitivity results. Isopleths of the deviations in maximum ozone concentrations are presented. It should be noted that maximum ozone concentrations may occur at the same location at different times. This isopleth presentation provides useful information about the magnitude and location of changes in ozone concentrations.

---

* The coastal stations that are not included in this analysis are Costa Mesa, El Toro, Laguna Beach, Long Beach, Los Alamitos, Redondo Beach, and West Los Angeles.

9. Ozone Profiles at Air Quality Monitoring Stations

Comparison of calculated and observed ozone concentrations at various air quality monitoring stations is another useful measure of model sensitivity; accordingly, the time-varying ozone concentrations at the monitoring stations are compared for the sensitivity and base cases. This provides information on the temporal variation and magnitude of the perturbations in ozone concentrations at various locations throughout the basin.

To keep the display of simulation results to a manageable level, 6 stations were selected from the 23 monitoring sites for detailed discussion of the sensitivity results. (Results for all stations are given in appendix A.) The six monitoring locations are:

> Anaheim
> La Habra
> Lynwood
> Pasadena
> San Bernardino
> Upland.

These stations were chosen because they typically present discernible deviations in ozone concentration levels between the sensitivity and base cases. Moreover, the stations are aligned with the wind flow trajectory that carries the photochemical plume eastward across the basin. Where indicated, additional monitoring stations that present interesting features in the diurnal ozone profiles will be discussed in some sensitivity studies.

10. Summary of Sensitivity Measures

These sensitivity measures are considered in the following discussion:

> Signed deviation, absolute deviation, temporal correlation, and spatial correlation of ozone levels for the following selected station locations and grid cells.

  - Monitoring station grid cells with ozone levels above 12 pphm.

  - Monitoring station grid cells with ozone levels above 20 pphm.
  - Grid cells with ozone levels above 12 pphm.
  - Grid cells with ozone levels above 20 pphm.
- Overall maximum ozone level predicted by the sensitivity case and the base case and the corresponding location of occurrence for:
  - Station statistics (i.e., statistics computed based on ozone predictions in grid cells containing monitoring stations).
  - Grid statistics (i.e., statistics based on ozone predictions in all ground level grid cells).
- Normalized differences and peak time lag of the ozone peak level for:
  - All 23 stations
  - 16 downwind stations.
- Dosages of the sensitivity case and of the base case, and the corresponding normalized difference.
- Isopleths of maximum ozone deviations between the sensitivity case and the base case.
- Diurnal ozone concentration profiles predicted by the sensitivity case and the corresponding base case at some monitoring stations.

Appendix A (bound separately) presents the detailed results for each simulation.

C. SUMMARY OF SENSITIVITY RESULTS

Major features of the sensitivity simulations are summarized in table V-1. Signed deviations, absolute deviations, temporal correlations, and spatial correlations are presented in tables V-2 through V-5 for ozone levels above 12 pphm and 20 pphm, for both station and grid statistics.

TABLE V-1. SUMMARY OF SENSITIVITY STUDIES

Sensitivity Study	Day of Simulation*	Change in Input Data
1	J	Reduced upper air meteorology
2	A	Reduced upper air meteorology
3	J	Reduced upper air and surface meteorology
4	A	Reduced upper air and surface meteorology
5-1	J	Reduced air quality
5-2	J	Reduced air quality and meteorology
6	J	Reduced upper air quality
7	A	Reduced upper air quality
8	J	Nitrogen oxides boundary conditions
9	J	Hydrocarbon initial condition
10	A	Hydrocarbon initial condition
11	J	Hydrocarbon speciation
12	A	Hydrocarbon speciation
13	J	Mobile sources--Older inventory
14	J	Mobile sources--Less detailed inventory
15	A	Mobile sources--Less detailed inventory
16	J	Point sources--Temporal resolution
17	J	Area sources--Spatial resolution
18	A	Area sources--Spatial resolution
19	J	Area sources--Temporal resolution
20	J	Larger grid size (10 km)
21	J	Reduction of grid cell layers--two-layer model
22	J	Reduction of grid cell layers--single-layer model

---

* J = 26 June 1974, A = 4 August 1975.

TABLE V-2. SENSITIVITY MEASURES FOR OZONE CONCENTRATIONS ABOVE 12 pphm

Station Statistics

Sensitivity Study	Signed Deviation	Absolute Deviation	Temporal Correlation	Spatial Correlation
1	-0.177	0.300	0.023	0.622
2	-0.094	0.164	0.912	0.638
3	-0.172	0.320	-0.183	0.310
4	-0.146	0.189	0.799	0.648
5-1	0.000	0.028	0.990	0.986
5-2	-0.208	0.329	-0.183	0.313
6	0.105	0.106	0.981	0.975
7	-0.073	0.073	0.980	0.953
8	-0.141	0.145	0.914	0.927
9	-0.344	0.344	0.869	0.800
10	-0.450	0.450	0.680	0.578
11	0.046	0.046	0.989	0.985
12	0.121	0.125	0.969	0.818
13	0.027	0.037	0.988	0.983
14	0.064	0.067	0.988	0.983
15	0.115	0.127	0.977	0.738
16	-0.034	0.035	0.990	0.990
17	0.054	0.055	0.990	0.982
18	0.118	0.127	0.977	0.789
19	-0.005	0.007	0.990	0.990
20	0.148	0.193	0.848	0.634
21	-0.007	0.053	0.975	0.972
22	-0.250	0.377	0.787	0.347

TABLE V-3. SENSITIVITY MEASURES FOR OZONE CONCENTRATIONS ABOVE 20 pphm

Station Statistics

Sensitivity Study	Signed Deviation	Absolute Deviation	Temporal Correlation	Spatial Correlation
1	-0.342	0.366	-0.009	-0.133
2	-0.034	0.082	0.768	0.974
3	-0.367	0.389	-0.455	-0.185
4	-0.073	0.104	0.682	0.937
5-1	0.003	0.022	0.990	0.986
5-2	-0.390	0.402	-0.402	-0.184
6	0.110	0.110	0.983	0.959
7	-0.048	0.049	0.982	0.990
8	-0.177	0.181	0.933	0.613
9	-0.384	0.384	0.984	0.433
10	-0.486	0.486	0.280	0.990
11	0.035	0.035	0.990	0.989
12	0.085	0.085	0.961	0.940
13	0.022	0.029	0.987	0.969
14	0.075	0.075	0.989	0.973
15	0.082	0.084	0.955	0.922
16	-0.034	0.035	0.990	0.989
17	0.057	0.057	0.990	0.979
18	0.103	0.103	0.963	0.946
19	-0.005	0.021	0.990	0.990
20	-0.037	0.087	0.982	0.721
21	0.006	0.033	0.989	0.982
22	-0.235	0.332	0.790	0.526

TABLE V-4. SENSITIVITY MEASURES FOR OZONE CONCENTRATIONS ABOVE 12 pphm

	Grid Statistics			
Sensitivity Study	Signed Deviation	Absolute Deviation	Temporal Correlation	Spatial Correlation
1	-0.229	0.329	0.134	0.376
2	-0.046	0.168	0.878	0.396
3	-0.225	0.347	-0.034	0.247
4	-0.107	0.198	0.737	0.450
5-1	0.006	0.032	0.988	0.980
5-2	-0.253	0.355	-0.067	0.265
6	0.103	0.108	0.977	0.061
7	-0.071	0.073	0.981	0.965
8	-0.130	0.139	0.925	0.896
9	-0.317	0.318	0.807	0.800
10	-0.443	0.443	0.722	0.806
11	0.044	0.044	0.986	0.985
12	0.138	0.142	0.966	0.923
13	0.022	0.039	0.984	0.983
14	0.070	0.073	0.986	0.980
15	0.146	0.160	0.963	0.906
16	0.009	0.010	0.990	0.990
17	0.053	0.055	0.988	0.984
18	0.141	0.149	0.969	0.919
19	-0.008	0.010	0.990	0.990
20	0.100	0.200	0.831	0.397
21	-0.011	0.055	0.975	0.965
22	-0.067	0.207	0.776	0.701

TABLE V-5. SENSITIVITY MEASURES FOR OZONE CONCENTRATIONS ABOVE 20 pphm

Grid Statistics

Sensitivity Study	Signed Deviation	Absolute Deviation	Temporal Correlation	Spatial Correlation
1	-0.366	0.378	0.012	0.403
2	-0.108	0.171	0.321	0.439
3	-0.392	0.443	-0.293	0.331
4	-0.165	0.187	0.263	0.608
5-1	0.010	0.032	0.984	0.965
5-2	-0.416	0.459	-0.253	0.348
6	0.110	0.113	0.972	0.941
7	-0.053	0.057	0.950	0.970
8	-0.165	0.169	0.928	0.794
9	-0.357	0.357	0.893	0.735
10	-0.463	0.463	0.336	0.872
11	0.041	0.041	0.985	0.982
12	0.090	0.091	0.958	0.856
13	0.018	0.034	0.977	0.975
14	0.085	0.085	0.982	0.967
15	0.095	0.108	0.945	0.840
16	0.007	0.008	0.988	0.990
17	0.060	0.060	0.986	0.981
18	0.103	0.107	0.960	0.836
19	-0.009	0.012	0.989	0.990
20	-0.053	0.150	0.889	0.115
21	-0.000	0.041	0.981	0.949
22	-0.086	0.162	0.834	0.678

The signed and absolute deviations for ozone levels above 12 pphm are presented graphically as well, in figures V-2 and V-3, respectively (results of sensitivity run 22 were obtained late and are not included in these figures).

Simulations involving a reduction in the upper air meteorological data or a change in the magnitude of reactive hydrocarbon initial conditions result in the greatest signed and absolute deviations. Such deviations range from 15 to 45 percent. The change in grid size results in an absolute deviation of about 20 percent. Other perturbations in the input data result in absolute deviations of less than 15 percent.

Reductions in meteorological data result in the lowest temporal correlations, with values as low as 0.023 and -0.183 for the 26 June 1974 simulation. Changes in the hydrocarbon initial conditions also result in a low temporal correlation coefficient for the 4 August 1975 simulation. The same simulations (reduction in meteorological data, change in hydrocarbon initial conditions) also lead to low spatial correlation coefficients. The change in grid size results in a low spatial correlation (i.e., $\rho < 0.721$), as is to be expected since the spatial features of the model have been modified.

Table V-6 presents the ratios of the overall maximum ozone level of the sensitivity case to the base case, along with the location of occurrence for station and grid statistics. The largest differences in overall maximum ozone levels are obtained for the simulations that involve a reduction in meteorological data or a variation in reactive hydrocarbon initial conditions. The lowest ratio of overall maximum ozone levels is obtained with the station statistics for sensitivity case 10 (hydrocarbon initial conditions).

Maximum ozone statistics are presented in table V-7. The peak level normalized differences and peak time lags are shown for the cases where all 23 stations and 16 downwind stations are considered. The largest difference between these statistics is obtained for simulation 5-1, which involves a reduction in air quality data. This result suggests that, for the reduction in air quality data, the perturbation is larger at the coastal stations.

The dosages and the corresponding normalized differences are presented in table V-8. In general, dosage is a sensitive measure of model perturbation; it is the most sensitive for simulation 20, since the change in grid size affects the computation of the dosage noticeably. Other dosage-normalized differences appear to be consistent with the sensitivity measures presented above.

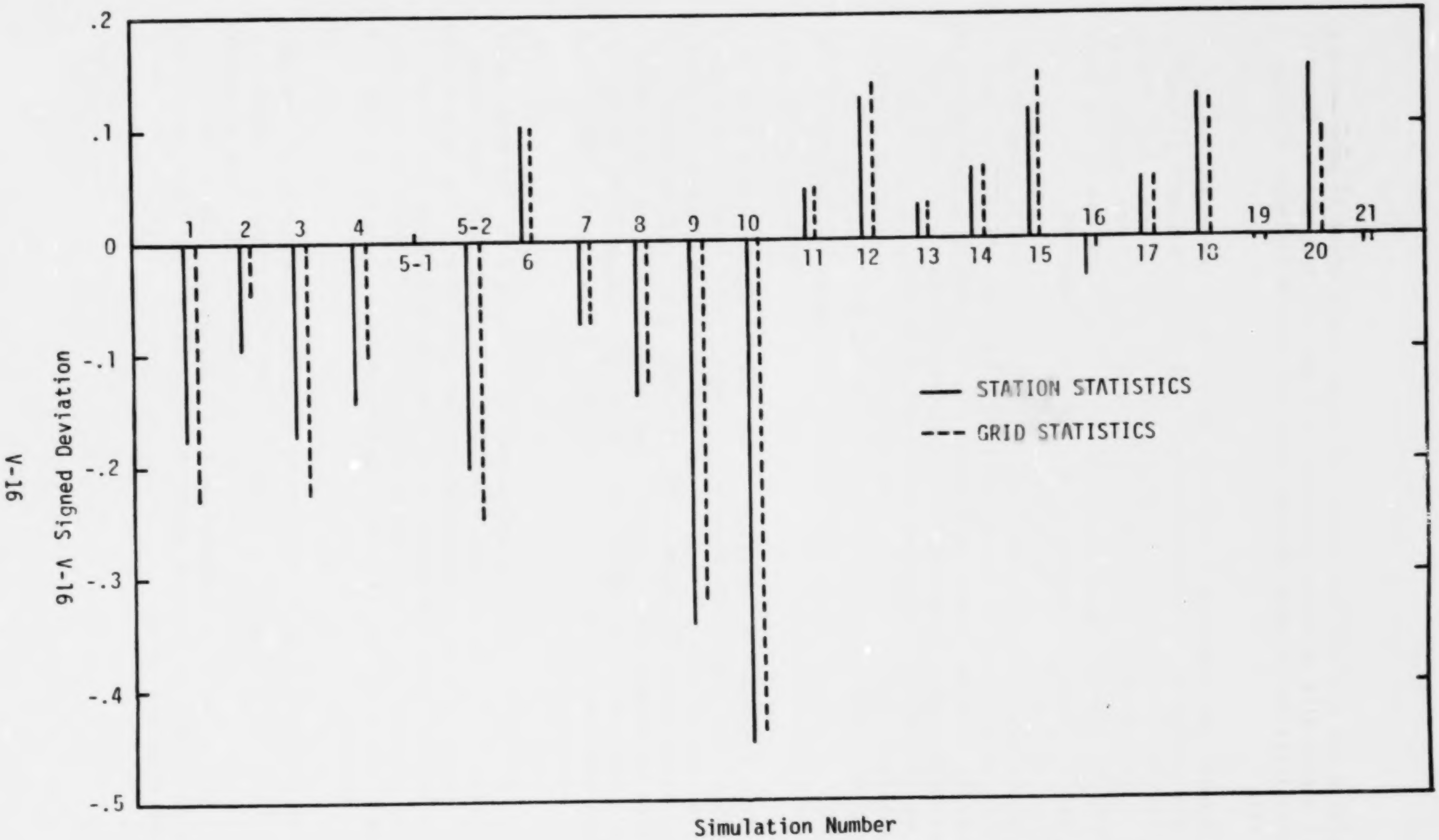


FIGURE V-2. SIGNED DEVIATIONS FOR OZONE CONCENTRATIONS ABOVE 12 PPHM

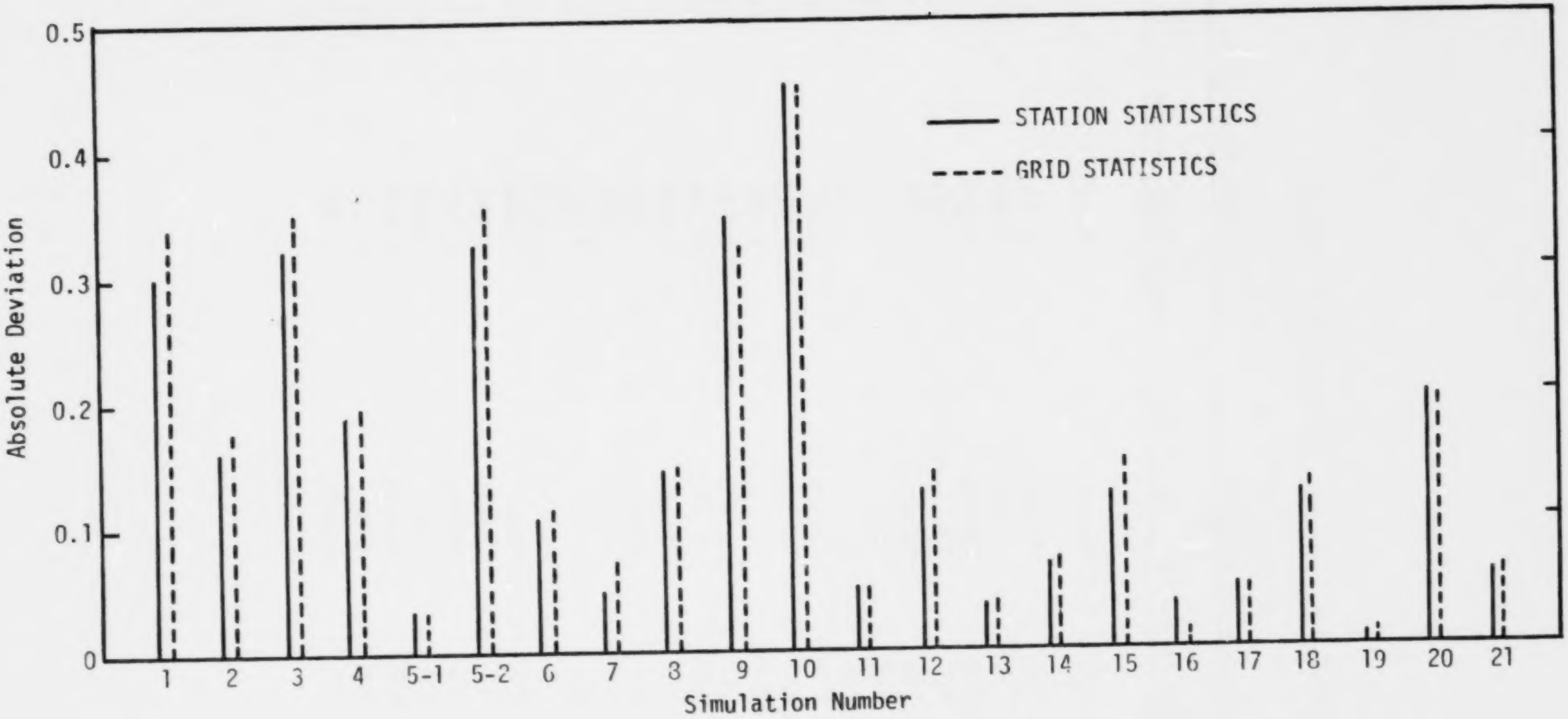


FIGURE V-3. ABSOLUTE DEVIATIONS FOR OZONE CONCENTRATIONS ABOVE 12 PPHM

TABLE V-6. OVERALL MAXIMUM OZONE LEVELS

	Station Statistics		Grid Statistics	
Simulation	Ratio Peak levels	Location	Ratio Peak Levels	Location (x,y)
Base case 26 June 1974	1.000	Azusa	1.000	33-14
Base case 4 August 1975	1.000	Pomona	1.000	30-16
1	0.718	Upland	0.814	30-16
2	1.038	Upland	0.839	11-16
3	0.710	San Bernardino	1.366	30-16
4	0.988	Upland	0.755	27-15
5-1	0.997	Azusa	0.982	19-15
5-2	0.683	Upland	1.251	30-16
6	1.109	Upland	1.122	19-15
7	0.927	Pomona	1.019	30-16
8	0.804	Azusa	1.002	33-14
9	0.632	Upland	0.832	33-14
10	0.523	Upland	0.670	30-16
11	1.023	Azusa	1.018	19-15
12	1.128	Pomona	1.003	30-16
13	1.020	Azusa	1.027	33-14
14	1.101	Azusa	1.056	19-15
15	1.139	Pomona	1.033	11-16
16	1.000	Azusa	0.996	33-14
17	1.067	Azusa	1.046	19-15
18	1.171	Pomona	0.997	11-16
19	0.988	Azusa	1.002	33-14
20	0.909	Pomona	0.854	20-14
21	1.019	Azusa	1.020	19-15
22	1.054	Azusa	0.947	14-15

TABLE V-7. MAXIMUM OZONE STATISTICS

Sensitivity Study	Peak Level Normalized Difference		Peak Time lag (hour)	
	23 Stations	16 Stations	23 Stations	16 Stations
1	0.253	0.258	1.77	1.87
2	0.128	0.133	0.57	0.50
3	0.249	0.260	2.17	2.37
4	0.143	0.121	0.82	0.81
5-1	0.061	0.031	0.52	0.00
5-2	0.294	0.266	2.25	1.93
6	0.162	0.144	0.56	0.12
7	0.057	0.069	0.04	0.06
8	0.188	0.187	0.87	0.75
9	0.320	0.354	0.96	0.56
10	0.394	0.476	1.00	0.81
11	0.114	0.092	0.30	0.44
12	0.303	0.313	0.25	0.39
13	0.028	0.029	0.35	0.44
14	0.070	0.082	0.13	0.19
15	0.243	0.214	0.12	0.26
16	0.029	0.020	0.13	0.19
17	0.064	0.063	0.13	0.19
18	0.254	0.241	0.12	0.26
19	0.008	0.001	0.00	0.00
20	0.155	0.112	0.54	0.75
21	0.050	0.043	0.22	0.25
22	--	Not Available	--	--

TABLE V-8. DOSAGES

Simulation	Dosage (km^2)	Normalized Difference
Base case: 26 June 1974	10,000	0.0
Base case: 4 August 1975	8,750	0.0
1	5,975	-0.4025
2	8,600	-0.0172
3	7,400	-0.2600
4	6,625	-0.2439
5-1	10,475	0.0475
5-2	7,125	-0.2875
6	14,525	0.4525
7	4,350	-0.5039
8	6,200	-0.3800
9	2,425	-0.7575
10	450	-0.9496
11	11,375	0.1375
12	13,975	0.5970
13	11,125	0.1375
14	12,450	0.1925
15	13,900	0.5900
16	10,125	0.0125
17	11,925	0.1925
18	13,750	0.5710
19	9,700	-0.0300
20	16,000	0.6000
21	10,425	0.0425
22	11,525	0.1525

Finally, table V-9 summarizes the results of the 22 sensitivity runs in terms of the various measures presented in this chapter. (These results are drawn from those presented in tables V-2 through V-8.)

D. SENSITIVITY RESULTS

In appendix A, results of the sensitivity studies summarized in this chapter are presented and discussed.* Model sensitivity is examined by means of various measures presented in this chapter (ozone level deviations, temporal and spatial correlations, comparisons of maximum ozone levels, and dosage). The largest perturbations caused by reduction in input data, manifested by changes in model input files, are presented. The sensitivity of model input variables, such as wind fields, mixing depths, and emission levels, to changes in available input data is analyzed to provide insight into how restrictions in data resources may affect model performance. Perturbations in ozone concentrations throughout the basin and at specific monitoring stations are also discussed.

In the following chapter, the sensitivity of the airshed model to perturbations in available input data is interpreted based on the results presented here and in appendix A.

---

* For each sensitivity study, corresponding reduction in input data is briefly introduced. A more detailed presentation of the procedure used to reduce the input data has been made in Section III.

TABLE V-9. SUMMARY OF AIRSHED MODEL SENSITIVITY RESULTS

Simulation Identification	Station Statistics for Ozone Concentrations above 12 pphm: Signed Deviation	Absolute Deviation	Temporal Correlation	Spatial Correlation	Station Statistics: Ratio of Peak Concentration Levels	Maximum Ozone Statistics for 16 Stations: Peak Level Normalized Difference	Peak Time Lag (hour)	Dosage Values*: Dosage (km^2)	Normalized Difference
1	-0.177	0.300	0.023	0.622	0.718	0.258	1.87	5,975	-0.4025
2	-0.094	0.164	0.912	0.638	1.038	0.133	0.50	8,600	-0.0172
3	-0.172	0.320	-0.183	0.310	0.710	0.260	2.37	7,400	-0.2600
4	-0.146	0.189	0.799	0.648	0.988	0.121	0.81	6,625	-0.2439
5-1	0.000	0.028	0.990	0.986	0.997	0.031	0.00	10,475	0.0475
5-2	-0.208	0.329	-0.183	0.313	0.683	0.266	1.93	7,125	-0.2875
6	0.105	0.106	0.981	0.975	1.109	0.144	0.12	14,525	0.4525
7	-0.073	0.073	0.980	0.953	0.927	0.069	0.06	4,350	-0.5039
8	-0.141	0.145	0.914	0.927	0.804	0.187	0.75	6,200	-0.3800
9	-0.344	0.344	0.869	0.800	0.632	0.354	0.56	2,425	-0.7575
10	-0.450	0.450	0.680	0.578	0.523	0.476	0.81	450	0.9496
11	0.046	0.046	0.989	0.985	1.023	0.092	0.44	11,375	0.1375
12	0.121	0.125	0.969	0.818	1.128	0.313	0.39	13,975	0.5970
13	0.027	0.037	0.988	0.983	1.020	0.029	0.44	11,125	0.1375
14	0.064	0.067	0.988	0.983	1.101	0.082	0.19	12,450	0.1925
15	0.115	0.127	0.977	0.738	1.139	0.214	0.26	13,900	0.5900
16	-0.034	0.035	0.990	0.990	1.000	0.020	0.19	10,125	0.0125
17	0.054	0.055	0.990	0.982	1.067	0.063	0.19	11,925	0.1925
18	0.118	0.127	0.977	0.789	1.171	0.241	0.26	13,750	0.5710
19	-0.005	0.007	0.990	0.990	0.988	0.001	0.00	9,700	-0.0300
20	0.148	0.193	0.848	0.634	0.909	0.112	0.75	16,000	0.6000
21	-0.007	0.053	0.975	0.972	1.019	0.043	0.25	10,425	0.0425
22	-0.250	0.377	0.787	0.347	1.054	--	--	11,525	0.1525

VI INTERPRETATION OF RESULTS

A. RANKING OF DATA NEEDS THROUGH SENSITIVITY-UNCERTAINTY ANALYSIS

A ranking of airshed model data needs must take into account the sensitivity of model predictions to input data, the cost of data acquisition, and uncertainties in the specification of model inputs. These factors will be considered in defining a sensitivity-uncertainty index that will provide a basis for ranking data needs.

1. Definition of a Sensitivity-Uncertainty Index.

The sensitivity of model predictions to the level of detail of input data is important in determining input data needs. Several sensitivity simulations have evaluated the effect on predicted ozone levels of perturbations in meteorological, air quality, and emission data. The results of these simulations provide quantitative information on model sensitivity. Several measures of model sensitivity, ΔJ, could be considered in developing a sensitivity-uncertainty index; here consideration is given to the absolute normalized mean deviation of ozone concentrations above 12 pphm and the absolute normalized deviation of the air-shed-wide peak ozone level. The former has been defined in equation (5-2).

The absolute normalized deviation of the peak ozone levels is computed from the peak ozone levels of the sensitivity case and base case, regardless of the grid of occurrence. Let ΔJ be the deviation of model predictions resulting from input data perturbations. In this case ΔJ is given in equation (6-1)

$$\Delta J = \left| \frac{C_{max,s} - C_{max,b}}{C_{max,b}} \right| \qquad (6\text{-}1)$$

where $C_{max,s}$ and $C_{max,b}$ are the air-shed-wide maximum ozone concentration in the sensitivity case and the base case, respectively.

Other measures, such as the deviation in dosage, might also be considered, but here we restrict our focus to the two measures given above.

In the following presentation subscript i refers to the meteorological conditions (26 June 1974 or 4 August 1975) and subscript j to the input data perturbed.

A sensitivity index is usually defined as the ratio of some model output deviation to the corresponding model input perturbation. If ΔI_j is the perturbation in the input data, the sensitivity index for the input data j and meteorological conditions i (26 June 1974 or 4 August 1975) is defined as follows:

$$\Gamma_{ij} = \frac{\Delta J_{ij}}{\Delta I_j} \quad . \tag{6-2}$$

This sensitivity index, Γ_{ij}, is a measure of the effect of a change in the model input data on the model predictions. In general, the deviation in the model predictions, ΔJ, may be expressed as follows:

$$\Delta J = \frac{\Delta \overline{C}}{\overline{C}} \quad , \tag{6-3}$$

where $\overline{C}$ is some model output measure such as the ozone level above 12 pphm or the overall maximum ozone level.

Model input perturbations differ in many cases, according to whether meteorological, air quality, or emission data were considered. To compare the various sensitivity studies, it is necessary to normalize these input perturbations. This may be done by introducing a cost index that relates the cost ΔK_j of acquisition of needed data (to raise the level of detail in input data j from the sensitivity case to the base case) to the change in input data:

$$\gamma_j = \frac{\Delta K_j}{\Delta I_j} \quad . \tag{6-4}$$

Next, the sensitivity measure is normalized with respect to a cost difference corresponding to the change in input data, because the ranking of data needs clearly must include such information. The normalized sensitivity index is then defined as follows:

$$S_{ij} = \frac{\Gamma_{ij}}{\gamma_j} \quad , \tag{6-5}$$

that is,

$$S_{ij} = \frac{\Delta J_{ij}}{\Delta K_j} \quad . \tag{6-6}$$

This sensitivity index is a measure of the effect of the cost (of improving the level of detail of input data j) on the model predictions for meteorological conditions i. It is, however, a "perturbation--specific" measure of model sensitivity. That is, the value of the sensitivity index has been evaluated for a specific perturbation. The classical sensitivity index, on the other hand, describes the sensitivity of the system to any perturbation and is a function of the perturbation level unless the sensitivity relationship is linear. As shown in figure VI-1, when the sensitivity relationship is nonlinear, the "perturbation-specific" and classical sensitivity indexes are identical only at the point where the former has been calculated.

Actually, the introduction of a cost variable probably makes the sensitivity relationship nonlinear, since a minimum threshold cost is necessary before an input data file may be improved. Before that threshold, the model is insensitive to the cost and the classical sensitivity index is zero; in that range, the "perturbation-specific" index is not representative of the model sensitivity. Therefore, when applying the results of these sensitivity studies to a different urban area, one should consider variations in the input data similar to those variations considered in this work. This will allow for the best use of the information obtained from this sensitivity analysis, which provides "perturbation-specific" sensitivity indexes. However, the general methodology employed here should be applicable, in principle, to other areas with differing characteristics.*

The need for increasing the detail of input data should be weighted by the accuracy available for corresponding model input variables. For instance, an updated mobile source inventory will provide detailed information on mobile source emissions of NO_x and NMHC. However, these emissions are uncertain by at least 15 and 25 percent, respectively; thus,

---

* It should be noted that it would not be appropriate to conduct the sensitivity analysis with respect to monitored data, since the detailed effects of the model input variables are interrelated and complex, and some misleading compensating phenomena could occur during a particular sensitivity simulation. For instance, the enlargement of the grid size gave better results (based upon few performance measures) than the base case when compared with experimental data.

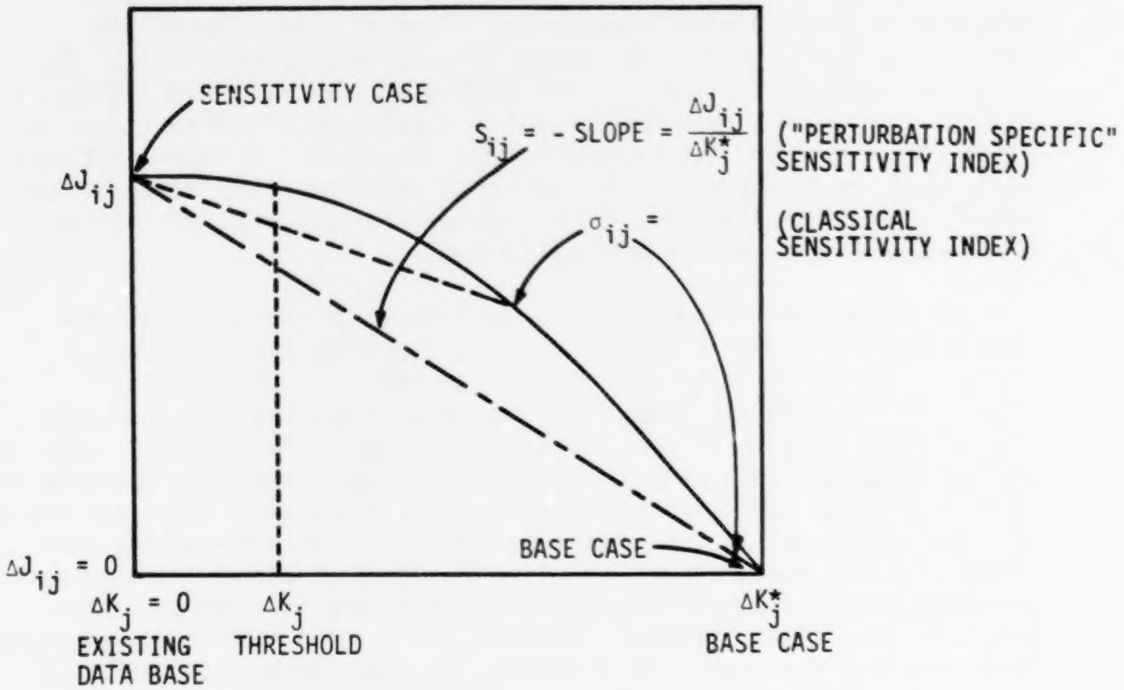

SENSITIVITY CASE
S_{ij} = - SLOPE = $\frac{\Delta J_{ij}}{\Delta K_j^*}$
("PERTURBATION SPECIFIC" SENSITIVITY INDEX)
σ_{ij} =
(CLASSICAL SENSITIVITY INDEX)
ΔJ_{ij}
ΔJ_{ij} = 0
BASE CASE
ΔK_j = 0
EXISTING DATA BASE
ΔK_j
THRESHOLD
ΔK_j^*
BASE CASE

FIGURE VI-1. COMPARISON OF "PERTURBATION-SPECIFIC" AND CLASSICAL SENSITIVITY INDEXES FOR A HYPOTHETICAL CASE

the available emission data are clearly limited. Since data for NO_x emissions are more reliable than data for RHC emissions, they would be of more use in improving the accuracy of model predictions.

Note that uncertainties in the model input variables as discussed here (see figure V-1) result from limitations in a mathematical, physical, or chemical procedure, such as the uncertainty in wind direction that exists in the computed wind field or the uncertainty in emission data for a detailed mobile source inventory. The uncertainty of model input variables due to the level of detail of the input data is not considered in the uncertainty index. It is taken into account in the sensitivity analysis: a change in input data introduces a change (or uncertainty) in the model input variables that leads to a perturbation in the model output. This uncertainty in the model input variables is therefore implicit in the sensitivity index, which could be rewritten as follows:

$$\Gamma_{ij} = \frac{\Delta J_{ij}}{\Delta P_{ij}} \cdot \frac{\Delta P_{ij}}{\Delta I_j} \quad , \qquad (6\text{-}7)$$

where ΔP_{ij} is the change in the model input variable P_{ij}, that is, the uncertainty in P_{ij} resulting from the level of detail of the input data.

The following uncertainty index is defined:

$$U_{ij} = \frac{\delta P_{ij}}{P_{ij}} \quad , \qquad (6\text{-}8)$$

where δP_{ij} is the uncertainty in P_{ij} resulting from limitations in a mathematical or physico-chemical representation of a process. It is defined for the input data j and the meteorological conditions i. The model input variable P_{ij} is the variable that most affects model performance during the sensitivity study defined by i and j. It may be different, depending on meteorological conditions. For example, a reduction in upper air meteorological data seems to modify the wind field for the 26 June 1974 simulation most, whereas the wind field for the 4 August 1975 simulation was not as affected; in contrast, for the August run, the mixing height is the most perturbed input variable.

The uncertainty index is defined as a dimensionless number. This allows consistency among uncertainties in input variables of a different nature (meteorological, chemical, or emission variables). Considering dimensionless perturbations and indexes is a common procedure in sensitivity-uncertainty analysis (Heinrich et al., 1977; Dougherty et al., 1979).

The larger the uncertainty in the input variable, the less need for additional input data. A very inaccurate parameter of the model lowers the need for improving the input data base to compute its value (which has an inherent uncertainty). Hence, we weight the sensitivity index S_{ij} by means of the uncertainty index U_{ij}, and a sensitivity-uncertainty index is defined as follows:

$$\Sigma_{ij} = \frac{S_{ij}}{U_{ij}}$$

$$\Sigma_{ij} = \frac{\Delta J_{ij}}{\Delta K_j} \cdot \frac{P_{ij}}{\delta P_{ij}} \quad . \qquad (6\text{-}9)$$

A large value of the index Σ_{ij} corresponds to a high ranking for data needs for the set of input data j under the meteorological conditions i for the Los Angeles urban airshed simulations.

2. Cost Estimates for Data Acquisition

Estimates of cost for acquiring data to upgrade a data base from one corresponding to the sensitivity case to that of the base case have been made and are listed in table VI-1 for the various sensitivity cases considered in this analysis. The assumptions made to derive these costs are also listed. Cost estimates are given as a range of values because of the uncertainty in data acquisition costs and, in some cases, because of the nature of the assumptions (which may define an upper and a lower bound).

These cost estimates should be seen as an illustration of the procedure for defining sensitivity-uncertainty indexes. Application of this analysis to a different urban area will undoubtedly require updating and modifying of these estimates.

3. Uncertainty Estimates for Model Input Variables

The chemical species concentrations are computed by solving a set of N continuity equations.

$$\frac{\partial c_i}{\partial t} + \nabla \underline{v}\, c_i = \nabla \underline{\underline{K}} \nabla c_i + R_i + S_i + L_i \qquad i = 1, \ldots, N \quad , \qquad (6\text{-}10)$$

TABLE VI-1. COST ESTIMATES FOR DATA ACQUISITION (FOR THE SOUTH COAST AIR BASIN)

Input Data	Improvement of Data	Assumptions	Cost Estimate
Upper air meteorology	Upper winds, mixing depth	8 week program (10-12 sampling days): 3 acoustic radars, 4 daily aircraft soundings, 4 daily pibals at 3 locations.	$ 50,000 - 100,000
Surface meteorology	Surface winds, temperature	Surface meteorological network expanded during eight week summer smog season. 6 meteorological stations.	20,000 - 30,000
Surface air quality	Ambient concentrations of NO_x, RHC, and O_3	Surface air quality network expanded during eight week summer smog season, 6 air quality monitoring stations (upper bound for permanent stations, lower bound for temporary stations).	75,000 - 150,000
Upper air quality	Upper air concentrations of NO_x, RHC and O_3	Eight week program (10-12 sampling days). Early morning, morning and noon soundings at upwind and major source regions. Nighttime sampling aloft over entire airshed.	50,000 - 125,000
IC for RHC	Relationship to relate RHC to THC	Reactive hydrocarbon monitoring for an eight week period during the summer smog season. Monitoring upwind and at major source region sites.	75,000 - 150,000
RHC speciation	Distribution of HC speciation according to source categories	Lower bound: Breakdown of sources into categories. Assume known HC speciation for each category. Upper bound: Literature review of hydrocarbon emissions speciation data. Limited source testing of major categories of emissions sources.	20,000 - 100,000
Mobile sources: Old inventory simulation	Update of an old inventory	Old inventory available. Acquisition of updated vehicle origin-destination data.	5,000 - 20,000
Mobile sources: Gas sales simulation	Creation of an inventory	Routines such as MOBILE 1 and EMFAC 5 available. Creation of vehicle origin-destination data files.	250,000 - 1,000,000
Point sources	More specific resolution of emissions rates	Characterization of day-specific emissions rates.	10,000 - 50,000
Area sources	Better spatial resolution	Characterization of spatial distribution of area sources.	50,000 - 100,000
Area sources	Better temporal resolution	Characterization of temporal profiles of all major stationary sources.	60,000 - 150,000

where

c_i = the concentration of species i,
$\underline{v}$ = the wind vector,
$\underline{\underline{K}}$ = the eddy diffusivity tensor,
R_i = the reaction rate,
S_i = the emission rates,
L_i = the removal rates,
t = the time,
N = the number of species considered in the model.

Initial conditions and boundary conditions are prescribed. Boundary conditions must be given at the boundaries of the airshed both at ground level and at the top of the mixing layer.

Input variables in this system of equations are the initial and boundary species concentrations, the wind vector, the eddy diffusivity tensor, the rate parameters (which depend on irradiation intensity and temperature), the emission and removal rates, and the depth of the mixing layer. For convectively mixed summertime conditions the eddy diffusivity is not a parameter to which the airshed model is highly sensitive. This may be compared with wind velocity variations that may substantially alter model predictions (Liu et al., 1976). Reaction and removal rate coefficients are not affected in the sensitivity studies considered here. Therefore, estimates for uncertainty are restricted to the following variables: wind velocity, species concentrations, emission rates, and mixing depth.

The importance of wind direction on model performance depends on the spatial distribution of emission sources. If the spatial distribution of emission rates were uniform throughout the airshed, wind direction would have little influence on the predicted air quality.* If, however, a strong gradient exists in the emission source distribution, the wind direction may be a major variable. According to average emission rate isopleths of the Los Angeles basin, an estimate of the gradient in emission sources corresponding to wind direction uncertainties* has been made.

---

* It would have some effect, however, because the airshed has a finite size, and the amount of precursor transported to a receptor will vary according to the travel from the boundaries.

Estimation of uncertainties in chemical species concentrations was restricted to nitrogen oxides and hydrocarbons, since these are the main precursors of photochemical smog. Uncertainties resulting from averaging these concentrations over each grid cell were not considered. Only uncertainties resulting from ambient monitoring were used. The uncertainty in the relationship between reactive hydrocarbons and total hydrocarbons was estimated from the experimental data used to derive the relationship.

The model input variables considered in this analysis are listed in table VI-2, along with the corresponding uncertainties.

4. Ranking of Data Needs

The sensitivity-uncertainty indexes were calculated as follows: For instance, the reduction in upper air meteorological data (set 1, j =1) for the 26 June 2974 (i=1) simulation leads to a relative absolute deviation in ozone levels above 12 pphm,

$$\Delta J_{11} = 0.33.$$

A cost estimate for data acquisition ΔK_1 = \$50,000, leads to

$$S_{11} = \frac{\Delta J_{11}}{\Delta K_1} = 6.6 \times 10^{-6}.$$

When the upper air meteorological data are reduced on 26 June 1974, the wind field is the most perturbed model input variable and the associated uncertainty range is 10 to 50 percent. For V_{11} = 10 percent, the following value of the sensitivity uncertainty index is obtained:

$$\Sigma_{11} = \frac{S_{11}}{V_{11}} = 6.6 \times 10^{-5}.$$

Similarly, a lower bound is obtained with cost and uncertainty estimates of \$100,000 and 50 percent, respectively: $\Sigma_{11} = 6.6 \times 10^{-6}$.

The results of the sensitivity-uncertainty analyses are listed in tables VI-3 and VI-4 for 12 sets of input data for the average deviation in ozone levels above 12 pphm and the deviation in air-shed-wide peak

---

[*] Uncertainties in wind direction and wind speed depend on wind speed.

TABLE VI-2. UNCERTAINTY ESTIMATES FOR MODEL INPUT VARIABLES

Input Variable	Uncertainty	Reference
Wind direction*	20°-50°	Tesche and Yocke (1978)
Emission sources	10-50%	Present Study
Wind velocity	0.5-1 m/sec	Tesche and Yocke (1978)
Mixing depth	50-70%	C. D. Unger (1976) P. B. Russell and E. E. Uthe (1978)
NO_x concentration	6%	Burton et al. (1976)
RHC concentration	20-60%	Burton et al. (1976)
RHC = function (THC)	25-65%	Present study
NO_x emissions	10-20%	Present study
HC emissions	20-30%	Present study

---

* Emission sources uncertainty is related to wind direction uncertainty.

TABLE VI-3. SENSITIVITY-UNCERTAINTY INDEXES: SENSITIVITY TO OZONE LEVELS ABOVE 12 pphm

Input Data	Ozone Level Deviation[*]	Cost for Data Acquisition (dollars)	Model Input Component Affected	Uncertainty	Sensitivity-Uncertainty Index[†] x 10^6 ($\$^{-1}$)
Upper air meteorology	0.33 (J) 0.17 (A)	50,000-100,000	Wind field (direction) Emission source Mixing depth	25°-50° 10-50% 50-70%	6.6 - (21) - 66 2.4 - (4.1) - 6.8
Surface meteorology	0.007 (J) 0.06 (A)	20,000-30,000	Wind field Emission sources	10-50%	0.5 - (1.3) - 3.5 4 - (11) - 30
Surface air quality	0.03 (J)	75,000-150,000	NO_x, RHC	20-60%	0.33 - (0.82) - 2
Upper air quality	0.11 (J)	50,000-125,000	NO_x, RHC	20-60%	1.5 - (4.1) - 11
IC for HC	0.32 (J) 0.43 (A)	75,000-150,000	RHC	20-60%	3.6 - (8.7) - 21.3 4.8 - (11.7) - 28.6
HC speciation	0.045 (J) 0.14 (A)	20,000-100,000	RHC	20-60%	0.75 - (2.9) - 11.25 2.3 - (9.0) - 35
Mobile source updating inventory	0.04 (J)	5,000-20,000	NO_x, RHC emissions	10-30%	6.7 - (23) - 80
Mobile source gas sales	0.07 (J) 0.16 (A)	250,000-1,000,000	NO_x, RHC emissions	10-30%	0.23 - (0.8) - 2.8 0.53 - (1.8) - 6.4
Point sources	0.01 (J)	10,000-50,000	NO_x, RHC emissions	10-30%	0.67 - (2.6) - 10
Area sources spatial resolution	0.055 (J) 0.14 (A)	50,000-100,000	NO_x, RHC emissions	10-30%	1.8 - (4.5) - 11 4.6 - (11.5) - 28
Area sources temporal resolution	.01 (J)	60,000-150,000	NO_x, RHC emissions	10-30%	0.22 - (0.6) - 1.7

---

[*] J and A refer to the 26 June 1974 and 4 August 1975 simulations, respectively.

[†] Lower bound, geometric mean value, upper bound.

TABLE VI-4. SENSITIVITY-UNCERTAINTY INDEXES: SENSITIVITY TO AIRSHED-WIDE PEAK OZONE LEVEL

Input Data	Basinwide Peak Ozone Level Deviation[*‡]	Sensitivity-Uncertainty Index[§] x 10^6
Upper air meteorology	0.188 (J)	3.76 - (11.9) - 37.6
	0.159 (A)	2.3 - (3.8) - 6.4
Surface meteorology	0.365 (J)	24.3 - (66.6) - 182.5
	0.243 (A)	16.2 - (44.4) - 121.5
Surface air quality	0.018 (J)	0.2 - (0.5) - 1.2
Upper air quality	0.121 (J)	1.6 - (4.4) - 12.1
IC for HC	0.167 (J)	1.9 - (4.6) - 11.1
	0.330 (A)	3.7 - (9.0) - 22.0
HC speciation	0.018 (J)	0.3 - (1.15) - 4.5
	0.005 (A)	0.08 - (0.3) - 1.25
Mobile source updating inventory	0.026 (J)	4.3 - (15.0) - 52
Mobile source gas sales	0.056 (J)	0.2 - (0.7) - 2.2
	0.035 (A)	0.1 - (0.4) - 1.4
Point sources	0.005 (J)	0.3 - (1.2) - 5.0
Area sources spatial resolution	0.045 (J)	0.15 - (1.15) - 9.0
	0.003 (A)	0.1 - (0.25) - 0.6
Area sources temporal resollution	0.0(J)	0.0

---

[*] The basinwide peak ozone level deviation is the absolute relative difference between the sensitivity and base case basinwide peak ozone levels.

[‡] J and A refer to the 26 June 1974 and 4 August 1975 simulations, respectively.

[§] Lower bound, geometric mean value, upper bound.

ozone levels, respectively. Rankings of data needs based on the sensitivity-uncertainty indexes for the 26 June 1974 and 4 August 1975 simulations are shown in figures VI-2 through VI-5. Note that for the average deviation in ozone levels above 12 pphm, only the ranking of meteorological input data (upper-air meteorology and surface meteorology) is affected by the meteorological conditions. Although the values of the sensitivity-uncertainty indexes of the other input data (air quality, chemistry, emissions) vary according to the meteorological conditions, their relative ranking is not affected. This suggests that meteorological conditions should be taken into account primarily when considering the need for meteorological input data.

Consider the 26 June 1974 sensitivity results for the average deviations in ozone levels above 12 pphm. Updating a mobile source inventory is ranked the highest. This is because of the relatively low cost involved and because of the reasonable accuracy of mobile source emission rates. Upper-air meteorology may be important; this depends, however, upon the meteorological conditions. The initial conditions for reactive hydrocarbons are also ranked high, despite the relatively high cost of a large monitoring program. This, however, has been shown to be a key input to the urban airshed model. Model predictions may be greatly affected by variations in hydrocarbon initial and boundary conditions. Reactive hydrocarbons, which are a necessary precursor of photochemical smog, are difficult to measure accurately in ambient air; they constitute a small amount of the total mass of hydrocarbons, which is composed primarily of methane. They are usually determined via the relationship of their concentrations to nitrogen oxides or total hydrocarbon concentrations. However, there is considerable uncertainty involved in these empirical formulas.

Spatial resolution of area sources, upper-air pollutant concentration data, detailed hydrocarbon speciation, and a detailed point-source inventory are generally of similar importance in model performance. The importance of surface meteorology is of the same order, though it varies with meteorological conditions.

Surface air quality is ranked relatively low. This is because we assumed a 10-station network in the Los Angeles basin for the sensitivity study. The sensitivity relationship is probably nonlinear and, if only 2 or 3 stations had been considered, the importance of surface air quality probably would have been higher. This underscores the caveat that the results of this sensitivity-uncertainty analysis must be considered in light of the assumptions made in this particular sensitivity study. It would be misleading to assume that surface air quality data have little effect on model predictions in other applications.

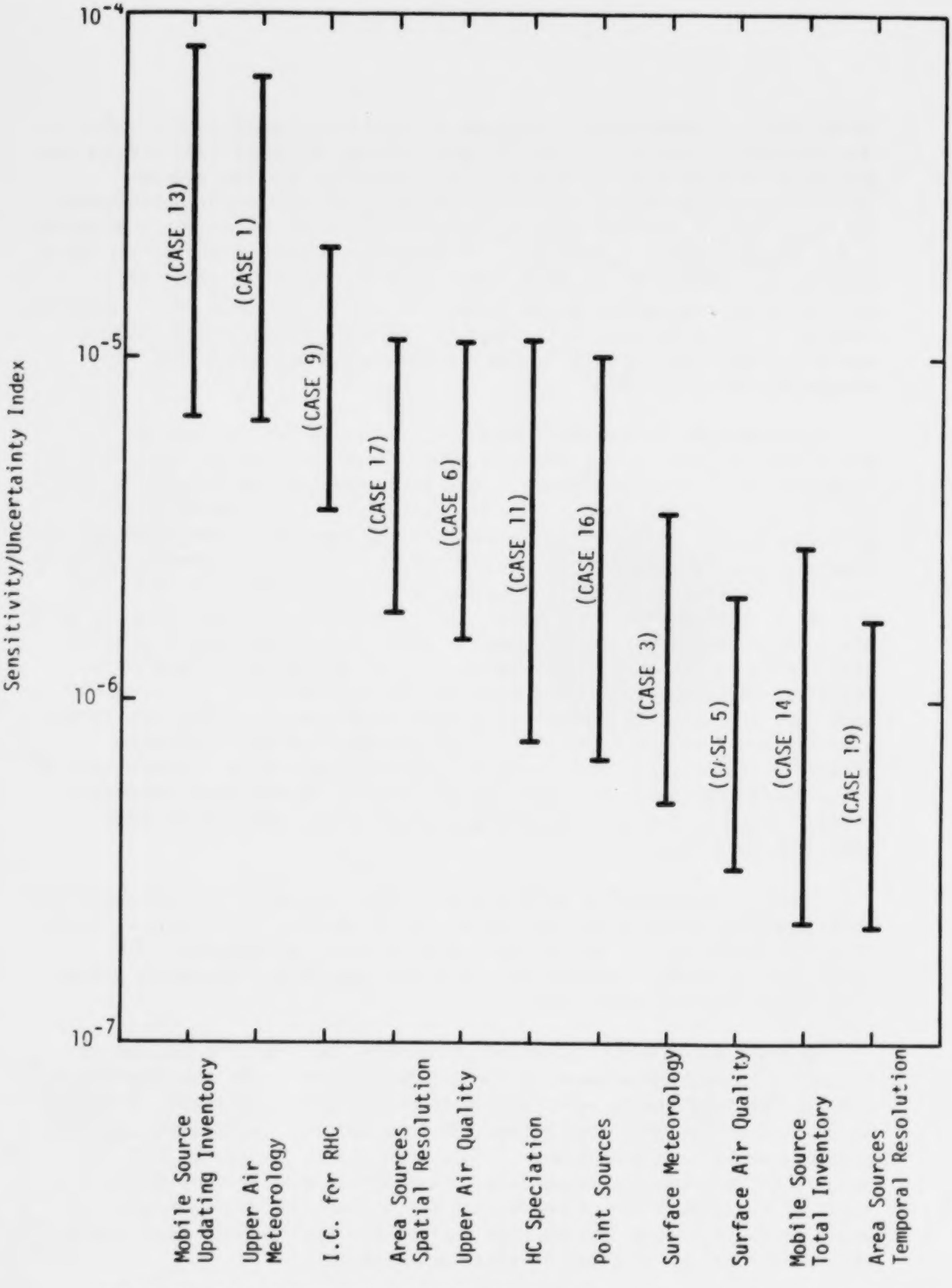


FIGURE VI-2. SENSITIVITY-UNCERTAINTY INDEXES--SIMULATION OF 26 JUNE 1974: SENSITIVITY TO OZONE LEVELS ABOVE 12 pphm

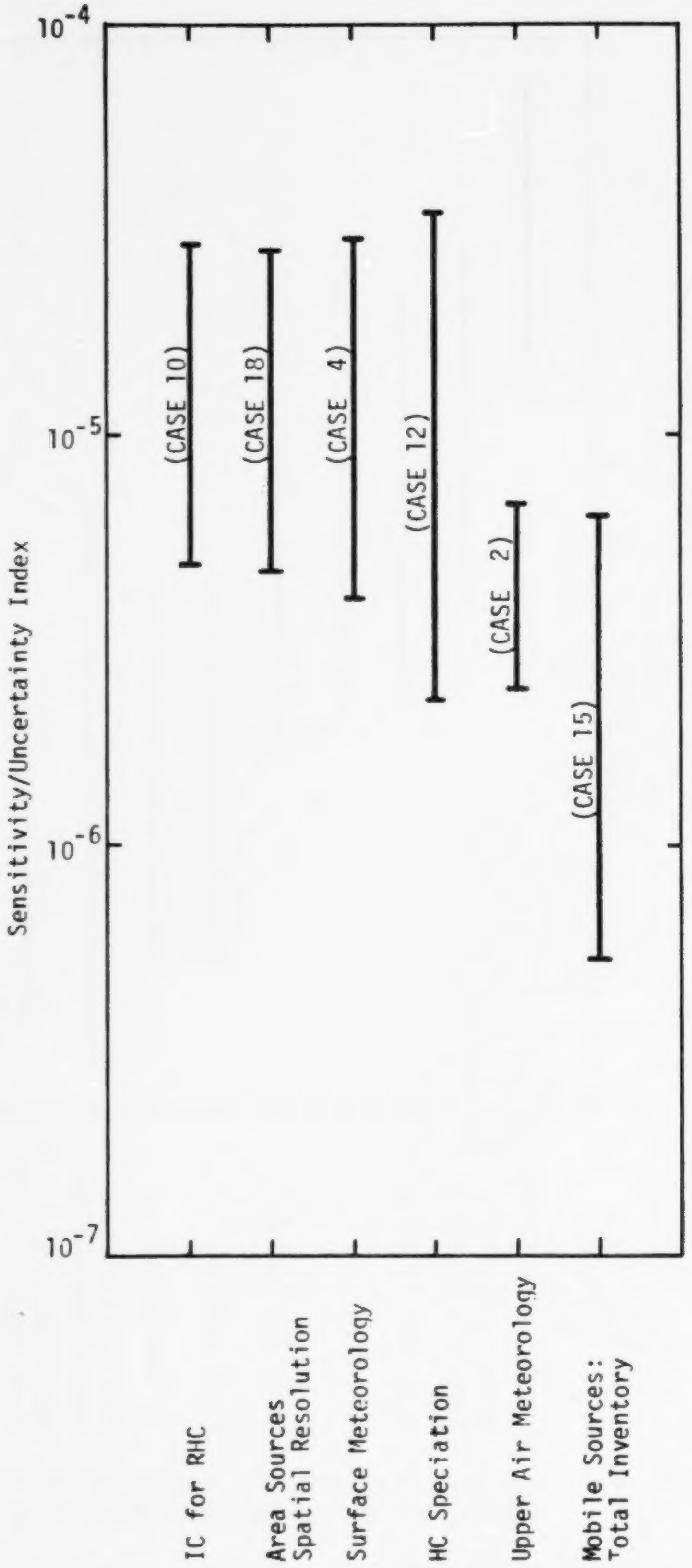


FIGURE VI-3. SENSITIVITY-UNCERTAINTY INDEXES--SIMULATION OF 4 AUGUST 1975: SENSITIVITY TO OZONE LEVELS ABOVE 12 pphm

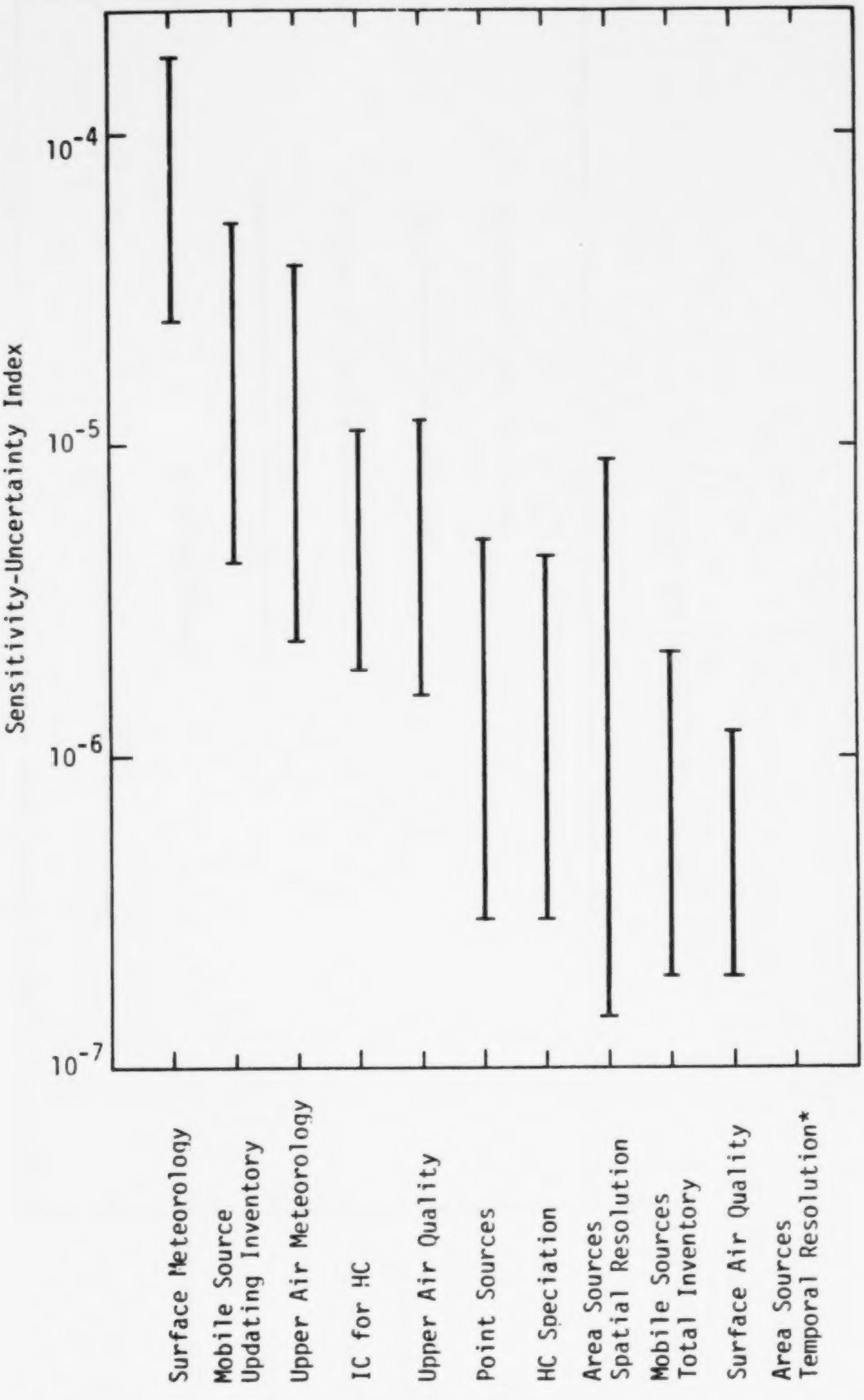


FIGURE VI-4. SENSITIVITY-UNCERTAINTY INDEXES--SIMULATION OF 26 JUNE 1974: SENSITIVITY TO AIR-SHED-WIDE PEAK OZONE LEVEL

* Sensitivity-uncertainty index is zero

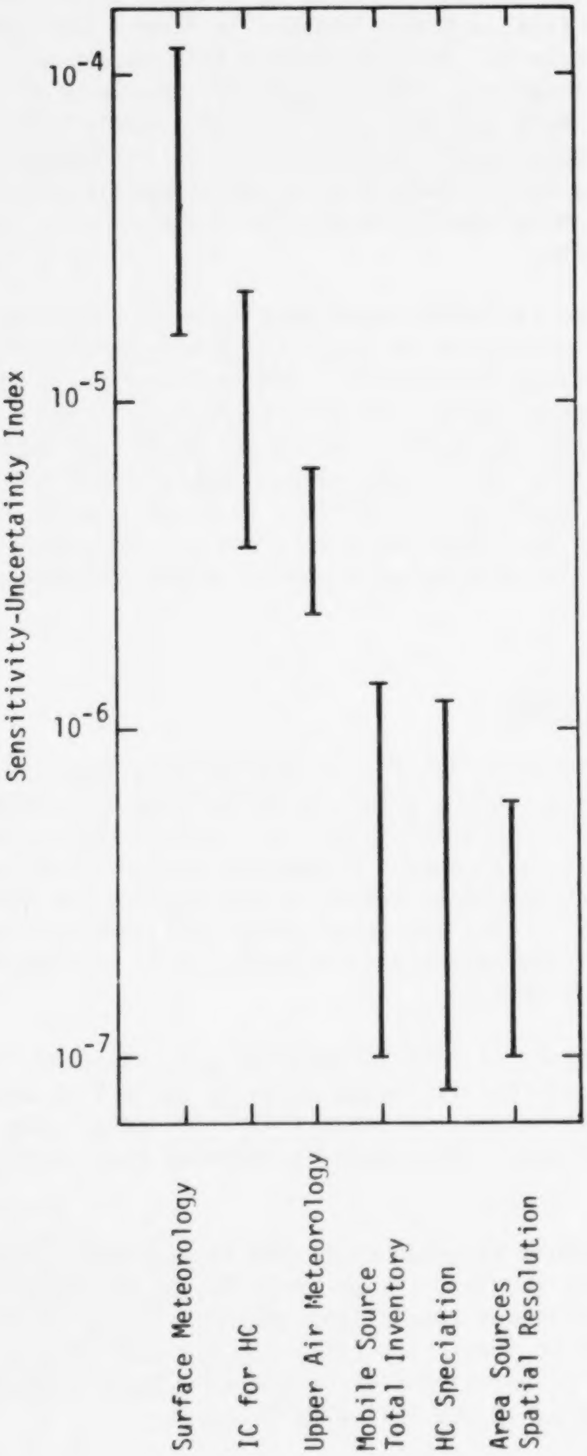


FIGURE VI-5. SENSITIVITY-UNCERTAINTY INDEXES--SIMULATION OF 4 AUGUST 1975: SENSITIVITY TO AIR-SHED-WIDE PEAK OZONE LEVEL

A detailed specification of transportation patterns is a rather expensive task, and this results in a very low ranking of the need for such input data. One may compare this result with the simple updating of the same inventory. This shows the importance of the assumptions that have been made and indicates that data needs will vary according to the existing data base. Quantification of the temporal distribution of stationary source emissions is rather expensive to obtain and has little effect on model performance. This results in a low sensitivity-uncertainty index.

If the basinwide ozone peak level, instead of the average ozone level, is considered as sensitivity measure, the ranking of the input data needs is slightly modified. The main changes occur for the surface meteorological data: The peak ozone levels appear to be more sensitive than the average ozone level to the number of meteorological stations at ground level, and surface meteorological data are ranked the highest for both simulation days (see table VI-4 and figures VI-4 and VI-5). The ranking of the other input data are modified within the uncertainties of the sensitivity-uncertainty index values and no other notable changes are observed.

5. Conclusions

The sensitivity-uncertainty analysis that has been presented should be seen as a procedure for defining input data needs, and the results obtained in this section should be considered as an illustration of the method. It is strongly recommended that revised costs and uncertainties specific to the photochemical model application area be estimated.. These city-specific cost and uncertainty estimates can then be used with the results of the sensitivity studies (ΔJ's) to compute the sensitivity-uncertainty indexes.

The types of results obtained are presented in figure VI-6. The hypothetical cost estimates shown in table VI-5 were used along with the uncertainty indexes of table VI-2. The priorities for acquisition of additional input data could be inferred from such information on data needs.

It should be noted that some limitations exist in the results of the sensitivity studies, though their extent may be difficult to estimate. These limitations result from the specificity of the Los Angeles basin, the length of simulation time, and the model characteristics (such as the treatment of point sources or the wind field submodel). These issues will be considered in the following section.

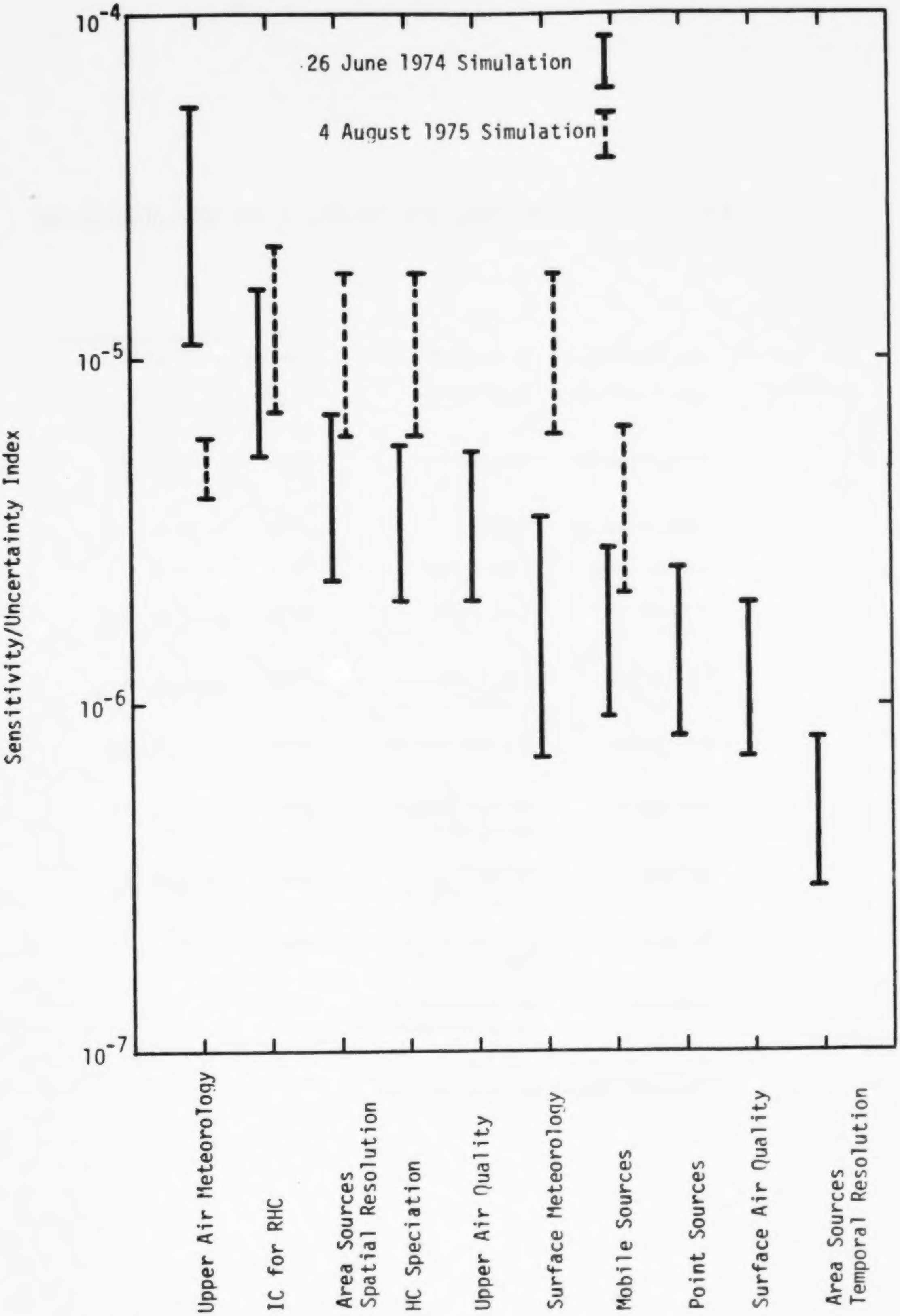


FIGURE VI-6. HYPOTHETICAL CASE OF NEEDS FOR DATA ACQUISITION

TABLE VI-5. HYPOTHETICAL COST ESTIMATES FOR DATA ACQUISITION

Input Data	Assumptions	Cost Estimate (dollars)	Sensitivity-uncertainty Index x 10^6*
Upper air meteorology	8 week program	60,000	11 - (25) - 55.0 4.0 - (4.8) - 5.7
Surface meteorology	6 stations	20,000	0.7 - (1.6) - 3.5 6 - (10) - 17.0
Surface air quality	6 stations	75,000	0.7 - (1.2) - 3.0
Upper air quality	8 week program	100,000	1.8 - (3.2) - 5.5
IC for RHC	8 week program	100,000	5.3 - (9.2) - 16.0 7.1 - (12.3) - 21.5
RHC Speciation	Known HC speciation by categories	40,000	1.9 - (3.2) - 5.6 5.9 - (10.0) - 17.4
Mobile sources	No inventory available	250,000	0.9 - (1.6) - 2.8 2.1 - (3.7) - 6.4
Point sources	Characterization of emission rates	40,000	0.8 - (1.4) - 2.5
Area Sources	Characterization of spatial distribution	80,000	2.3 - (4.0) - 6.9 5.9 - (10) - 17.6
Area sources	Characterization of temporal distribution	120,000	0.3 - (0.5) - 0.8
SO_2 Emissions	Data for point source inventory available	20,000	

---

* Lower bound, geometric mean value, upper bound.

B. GENERALIZATION OF RESULTS

The sensitivity studies of the airshed model have been carried out based on one-day simulations of the Los Angeles basin. Because these results may be of general interest to the photochemical modeling community, it is helpful to define the specific features of the simulations and to evaluate the limitations of the sensitivity results. Then it will then be possible to extend the conclusions of this study and to outline general recommendations for the application of the study results.

1. Specific Attributes of the Simulations

The results of the sensitivity simulations depend on the perturbations introduced in the input data, the area chosen for the study, the atmospheric conditions, and the modeling conditions (e.g., type of wind model and length of simulation time). Limitations of the analyses have been discussed in the previous section and the effect of aerometric conditions has been considered in several sensitivity runs carried out under different meteorological and air quality conditions. At this point it is appropriate to consider the effect on sensitivity results both of the specific attributes of the Los Angeles basin and of the modeling conditions.

a. Specific Attributes of the Los Angeles Basin

The meteorology of the western coast of the United States is strongly affected by the presence of the North Pacific anticyclone. The existence of a subsidence inversion layer above the marine layer, resulting from the adiabatic warming of descending air, leads to the trapping of air pollutants under an inversion layer that lies typically at an elevation of about 500 to 1000 meters. The presence of a temporally and spatially varying mixing height in the Los Angeles basin affects the generality of the model sensitivity results, since it probably enhances the need for upper-air meteorological and upper air quality data.

The Los Angeles basin, because of its location, is also characterized by a coastal-desert meteorology. The land-sea-breeze effect is enhanced by the presence of warm inland areas such as the Mojave desert, and this recirculating meteorology largely determines the atmospheric transport and dispersion of pollutants in the basin.

It should be noted that this land-sea-breeze effect is observed in many locations. It has been suggested, for instance, that air pollution episodes observed in Milwaukee are induced by a land-lake-breeze effect

(Lyons and Cole, 1976); experimental and modeling studies in the Tampa/St. Petersburg area have retrieved the main features of the land-sea-breeze effect (Liu et al., 1979).

The location of the Los Angeles basin is such that pollution observed is a result of local emission of primary pollutants. This may not be the case for other urban areas where pollution episodes may be, in part, the result of pollutant emissions that occurred far upwind. Usually, the importance of the boundary conditions depends on meteorological conditions. For instance, a steady westerly flow can occur in the Los Angeles basin for several days (e.g., the air pollution episode of 4 August 1975). In this instance, the boundary conditions will have little importance, since the air flowing into the airshed should have typical background concentrations of pollutants. During the land-sea-breeze regimes, however, polluted air is carried offshore during the evening and the aged urban air mass is carried inland the next morning. In this case the definition of the western boundary conditions of the airshed is important. For other urban areas, various meteorological conditions can also affect the importance of air quality data at the boundaries of the airshed.

The topography of the Los Angeles basin is characterized by complex terrain. This feature will be discussed in more detail in the following section, since it determines the selection of the wind model. In general, the complex terrain increases the need for meteorological data.

The air quality network in the Los Angeles basin is dense and widespread; such a rich air quality data base is not available in other urban areas. This has been emphasized in the previous section on the ranking of data needs, and it should be taken into account in determining the importance of air quality data from these sensitivity studies.

Primary pollutants in the Los Angeles basin are emitted mainly from mobile sources, which comprise about 44 percent of the reactive hydrocarbon emissions and 50 percent of the total nitrogen oxides emissions. This should be considered in the generalization of those sensitivity studies that involve modifications of the emission inventories.

b. Modeling Conditions

The modeling conditions affect the performance and sensitivity of the model. Some issues, such as the grid size or the number of grid layers, have been addressed in the sensitivity analyses (sensitivity cases 20 and 22). Here some aspects of modeling conditions that may affect the generality of the results are considered.

The wind model is an important component of the urban airshed model, and the use of various wind models (interpolated wind model, two-dimensional, and three-dimensional wind models) has been investigated (Reynolds et al., 1979). In this study, a three-dimensional wind model developed for complex terrain was used (Yocke et al., 1979). The wind field is obtained from the solution of a parameterized three-dimensional boundary value problem. Therefore, this wind field preparation procedure is more sensitive to atmospheric data at the boundaries of the airshed and at the top of the mixing layer than are other procedures based, for instance, on interpolation by inverse weighting.

The mixing height and eddy diffusivities depend on time and location. The computational procedures used to determine these atmospheric parameters have been presented elsewhere (Reynolds et al., 1979; Killus et al., 1980). Although other procedures could be considered, the sensitivity of the air quality model to computational procedures for the mixing height and eddy diffusivities has not been investigated.

The limitation of model sensitivity results for emission inventories stems primarily from the averaging of emissions over the grid cells. Although an attempt to model microscale phenomena for line sources has been made (Reynolds et al., 1979), it was not used in these simulations. The scavenging of ambient ozone by nitrogen oxides emitted from power plants and automobile exhausts is a well-known phenomenon in air pollution. The present treatment of these sources in the urban airshed model may obscure, in part, the need for greater accuracy in the corresponding emission inventories and, possibly, in other parameters (e.g., a finer resolution of atmospheric dispersion.

The length of simulation time has some effect on the sensitivity results. For example, initial conditions at 0400 may represent a sizeable fraction of the total precursor emission burden between 0400 and the time of the ozone peak (1200 to 1400). If uncertainties exist about the boundary conditions of the pollutant concentrations, these uncertainties may affect the model sensitivity more than the uncertainties in emission levels will affect it. However, if a multiple-day simulation were performed, the influence of initial conditions on ozone levels would be markedly reduced (particularly for land-sea-breeze regimes), and the emission inventories would have a dominant role, since they would represent most of the precursor mass. Clearly, there is a need for analyzing the sensitivity of the urban airshed model for a multiple-day simulation, to assess the importance of air quality data and emission inventories to model predictions under such conditions.

2. Limitations of the Sensitivity Results

Limitations of the sensitivity results that derive from the specific attributes of the Los Angeles basin and the length of simulation time are presented in table VI-6. The importance of the air quality data and emission inventories to model predictions should be strongly affected by the length of simulation time. For the reasons mentioned above, when the simulation is extended from a single-day to a multiple-day simulation, the importance of air quality data will decrease, whereas the effect of emission inventories on model predictions will increase.

In sensitivity cases 1 and 2, the number of upper air meteorological stations was reduced. The existence in the Los Angeles basin of a strong inversion that varies with time and location enhances the need for upper air meteorological data. Although the need for upper air meteorological data in other urban areas depends on local meteorology, air pollution episodes are often associated with strong inversions. This would make the results of this study representative of a typical urban airshed simulation. Moreover, long-range transport of pollutants trapped in the rear part of an anticyclonic system above the inversion height is an important air pollution phenomenon in midwestern and eastern urban areas. To properly account for such phenomena, upper air meteorological and air quality data would be needed.

Clearly, the spatial variation of the mixing height in midwestern and eastern urban areas may be negligible in comparison, and one upper air meteorological station should generally provide the information desired. In these sensitivity studies, the ranking of meteorological input data depends on the meteorological conditions. The use of a three-dimensional wind model for the Los Angeles basin probably emphasizes the need for upper air data as well. This model depends on boundary values, and the reduction of input data at the upper boundary of the domain probably affects the accuracy of this model. A flat urban area would not require the use of this wind model; consequently, a simple interpolation wind model would suffice, reducing the need for upper air data.

In sensitivity cases 3 and 4, surface meteorological data and upper air meteorological data were reduced. The density of the surface meteorological network in the Los Angeles basin must be taken into account when interpreting the sensitivity results. In the previous section on the ranking of data needs, it was pointed out that the intensity of input data in the base case and in the sensitivity case was a major factor in the sensitivity results.

The same considerations apply to the reduction in air quality data intensity. The density of the aerometric network in the Los Angeles basin

TABLE VI-6. POSSIBLE LIMITATIONS OF THE SENSITIVITY RESULTS

Sensitivity Case	Input Data Affected	Model Input Variables Affected	Degree of Limitation Due to the Unique Attributes of the Los Angeles Basin*				Degree of Limitation Due to the Simulation Time (single-day simulation)*
			Meteorology	Complex Terrain	Dense Aerometric Network	Emissions Sources	
1,2	Upper air meteorology	Wind field Mixing depth Eddy diffusivities	S > Intense subsidence inversion > Coastal-desert meteorology	S > Wind model affected by boundary values	M	--	L
3,4	Surface and upper air meteorology	Wind field Mixing depth Eddy diffusivities	> Intense subsidence inversion > Coastal-desert meteorology	> Wind model affected by boundary values	S (highly non-linear sensitivity)	--	L
5	Surface air quality	Ground level initial and boundary conditions	L	--	S (highly non-linear sensitivity)	M > Local emissions only	S
6,7	Upper air quality	Initial and boundary conditions aloft	S > Intense subsidence inversion	--	M	M > Local emissions only	S
8	Boundary conditions	Nitrogen oxides boundary conditions	S > Land-sea breeze effect	--	M	M > Local emissions only	S
9,10	Reactive hydrocarbon concentrations	Ground level HC initial and boundary conditions	--	--	S > Detailed ambient RHC determinations performed by the ARB	L	S
11,12	Hydrocarbon speciation	Relative magnitude of olefins, paraffins, aromatics, carbonyls, and ethylene emissions	--	--	--	S > Refinery emissions affected	S
13,14,15	Mobile source inventory	Magnitude, location, and timing of motor vehicle emissions	--	--	--	L > Importance of mobile source emissions common to most urban areas	S
16	Point source inventory	Magnitude and temporal resolution of point source emissions	--	--	--	M > Small contribution of power plant source emissions	S
17,18	Area source inventory	Location of area source emissions	--	--	--	M > Little specificity of area sources	S
19	**Area source inventory**	Temporal resolution of area and point sources	--	--	--	M > Little specificity of stationary sources	S
20	Enlarged grid size	**Smoothing of gradients in aerometric and emission fields**	> Horizontal smoothing of meteorological field	S > Variable wind field	S > Horizontal smoothing of gradients	S > Horizontal smoothing of emission field	L
21	2-layer model	Vertical smoothing of concentration and meteorological field	S > Vertical smoothing of meteorological field	S > Variable wind field	L > Vertical smoothing of concentration field	M > Relatively few elevated sources	L

* S = strong influence, M = medium influence, L = little influence.

allows for a certain reduction in the number of air quality monitoring stations without inducing any major perturbations in the model predictions. It would be misleading, however, to conclude that the need for air quality input data is low, since the results of the sensitivity analysis depend on that intensity of data in both the sensitivity and the base case. That is, the results depend on the number of air quality stations used for the Los Angeles basin simulations.

The need for upper air quality data depends strongly on the existence and intensity of the inversion. It is probable that areas with weaker inversions would not be as sensitive to upper air quality data as is the Los Angeles basin. However, air pollution episodes are often induced by the existence of a strong inversion, and the results obtained for the Los Angeles basin should be fairly representative for the application of the urban airshed model to the simulation of urban air pollution episodes. The length of simulation time has a strong influence on sensitivities to both surface and upper air quality data. These input data determine the initial conditions for the continuity equations of the air quality model. Hence, their importance will decrease if the simulation is carried out over a longer period of time, and the model predictions will become more sensitive to emission data.

In sensitivity case 8, the boundary conditions for hydrocarbon and nitrogen oxides were modified. The length of simulation time would probably affect the importance of this perturbation on the model predictions, as mentioned above. The location of the airshed will affect the importance of the boundary conditions to some extent. For the Los Angeles basin, the model is particularly sensitive to the western boundary conditions during land-sea-breeze regimes. If the model is applied to an urban area affected by pollutants transported from emission sources located outside of the modeled area, it also will be sensitive to boundary conditions, and air quality data needs at the airshed boundaries may be higher than they are for the Los Angeles basin.

The importance of the initial conditions for reactive hydrocarbon concentrations has been investigated in sensitivity cases 9 and 10. The length of the simulation time strongly influences the importance of these initial conditions. It should be noted that the initial conditions for reactive hydrocarbon (RHC) concentrations correspond 1) to the estimation of reactive hydrocarbon concentrations from an estimated RHC/NO_x ratio (sensitivity case), and 2) to the more accurate estimation of the RHC concentrations from detailed atmospheric chemical data obtained by gas chromatography on reactive and total hydrocarbon (THC) levels (base case). Reactive hydrocarbon concentrations are difficult to obtain on a routine basis, since they constitute a minor part of the total hydrocarbon concentration, which consists mainly of methane. Consequently, concentrations of RHC, which are obtained from the difference between THC and methane concentration (rounded off to the nearest ppm), are highly uncertain. Reliable data may be obtained by gas chromatography, but such measurements are not presently made on a routine basis. Detailed ambient reactive hydrocarbon data are available for the Los Angeles basin, and this may not be the case for other urban areas. However, this factor is taken into account in the cost index for the sensitivity-uncertainty analysis performed to rank data needs.

In sensitivity cases 11 and 12, the hydrocarbon speciation was averaged over all source categories. Clearly, the effect on model predictions depends on the diverse nature of source categories. In the sensitivity results, it appeared that refinery emissions were perturbed, whereas the dominant mobile source emissions were not modified notably. Higher levels of olefins, aromatics, aldehydes, and ethylene were observed in the refinery emissions in the sensitivity case. These hydrocarbons are oxidized faster than paraffins in the atmosphere, and this leads to higher ozone levels downwind of the refineries. The importance of these input data could be evaluated by careful identification of the various source categories.

The length of simulation time (single-day or multiple-day) is clearly an important factor for this sensitivity study and the following ones, which involve emission inventories (cases 11 through 19).

The mobile source inventory was perturbed in sensitivity cases 13, 14, and 15. The importance of mobile source emissions in the Los Angeles basin may enhance the importance of the corresponding inventories. However, major urban areas present a similar pattern, and the sensitivity results should be fairly representative of large urban communities.

The point source emission inventory was perturbed in sensitivity case 16. The contribution of power plant emissions to the air pollution burden is small for the Los Angeles basin. The importance of this inventory may be slightly larger for an area with a large number of power plants and smelters.

In sensitivity cases 17 and 18, the area-source inventory was spatially perturbed. No specific features of the Los Angeles basin appear to limit the generality of these results notably.

The temporal resolution of area sources was perturbed in sensitivity case 19. Mobile sources are dominant in the Los Angeles basin; however, this sensitivity study should give a fair estimate of the sensitivity of an airshed model to area sources.

The enlargement of the grid size in sensitivity case 20 leads to an averaging of the model variables. The spatially variable meteorological fields of the Los Angeles basin are probably affected by this averaging procedure. However, this sensitivity study should be fairly representative of the general effect of the grid size on an urban airshed model.

The number of grid layers was reduced from four to two in sensitivity case 21. Specific features of the Los Angeles basin are the small number of elevated sources and the spatially varying mixing depths. These factors may affect the nature of the perturbations, but this sensitivity study probably gives a reasonable evaluation of the effect of the vertical smoothing of concentration and meteorological variables on air quality predictions

3. Generalization of the Results

It has been pointed out that the nature of the sensitivity analysis considered in this study is "perturbation-specific." Therefore, it is important to apply the sensitivity results within the context from which they have been derived. This is particularly apparent for the sensitivity case concerning surface air quality. The low ranking of the need for air quality data results from the density of the air quality monitoring stations in the Los Angeles basin; it does not imply that air quality data are a negligible component of the model.

In the presentation of the results of the sensitivity-uncertainty analyses, the use of cost estimates and uncertainty estimates specific to the urban area should be considered. This provides for optimal use of the methodology developed in this study, since these estimates may vary according to the urban area considered.

The limitations of the sensitivity analysis that result from the specific attributes of the Los Angeles basin and the modeling conditions have been discussed in the previous section. These limitations should be considered as a means for adjusting the results of the sensitivity studies to different urban areas. It is possible to use the information obtained for the Los Angeles area in another location by evaluating the specific attributes of both areas.

In conclusion, the results of the Los Angeles sensitivity studies using the airshed model may be applied in a general context as long as proper attention is paid to the features both of the basin and the urban area to be considered.

VII CONCLUSIONS

Because of the large amount of input data (meteorology, air quality, and emissions) necessary to run grid-based photochemical models, it is important to identify the input data that most affect model performance as well as the amount and quality of data that should be acquired before carrying out air quality simulations.

This study investigated the sensitivity of the SAI Urban Airshed Model to the amount (and consequently the quality) of the input data in order to develop a general procedure for specifying the meteorological, air quality, and emission data most needed to achieve good model performance. The approach to sensitivity analysis used in this study and the general procedure developed to define a ranking of input data needs is summarized in this chapter. The major results obtained from the simulations of the Los Angeles basin and possible limitations are then discussed. Finally, suggestions on further studies to improve the knowledge of photochemical model sensitivity to various levels of details in input information are presented.

A. GENERAL PROCEDURE FOR SENSITIVITY ANALYSIS

The sensitivity of the airshed model to the level of detail of the input data was analyzed for the Los Angeles basin under different meteorological and air quality conditions (26 June 1974 and 4 August 1975). The effects of perturbations in the meteorological, air quality, and emission data, as well as in the model structure (grid size and number of grid layers), are analyzed and discussed in detail in Chapter V. The generality of specific numerical results presented in this study may be limited owing to the specificity of the Los Angles basin and modeling conditions. However, the methodology used to analyze the sensitivity of the model to the amount of input data is generally applicable and constitutes a framework for the generalization of the results of this study and the definition of future sensitivity studies.

The type of sensitivity analysis considered in this study differs from classical sensitivity analysis: the concern is with the sensitivity of the ozone predictions of the model (which consists of the air quality model and several input generation models, such as the wind model, eddy diffusivity computations, and emission inventories) to the quality of input data supplied (e.g., number of meteorological or air quality stations, level of detail of emission inventories). Classical sensitivity analysis deals simply with the sensitivity of a mathematical model to its parameters and initial conditions (e.g., Liu et al., 1976).

It was shown in section V.A that the overall sensitivity of the model to input data can be considered as the combination of the sensitivity of the model to the appropriate input variables (e.g., wind velocity, mixing depth, emission term) and the sensitivity of these input variables to perturbations in the input data. For instance, the sensitivity of the ozone predictions to the number of surface meteorological stations was analyzed by considering the sensitivity of the wind field to the number of meteorological stations and the sensitivity of the ozone levels to the wind field (i.e., wind directions and wind velocities). It was thus possible to evaluate the dynamics of model sensitivity in detail and to assess the possible limitations of the results. Then, the analysis of the sensitivity results was used as a basis for outlining general recommendations for the acquisition of input data (see section VI.B).

Several measures involving ozone predictions were used to evaluate the sensitivity of the model to perturbations in the input data:

> The signed deviation

> Absolute deviation

> Temporal and spatial correlations (for ozone levels above 12 and 20 pphm and for both station and grid statistics).

> Overall maximum ozone level

> Maximum ozone statistics (peak levels and peak times).

> Dosage (area with ozone level above 20 pphm).

Whereas results obtained from these various sensitivity measures reflects the specific features of model sensitivity (e.g., the dosage is a very sensitive measure for the variation in grid size) there is also consistency among the results given by the measures.

The general methodology developed to evaluate the level of detail in input data required for grid-based photochemical modeling needs is based on the definition of a sensitivity/uncertainty index that provides a means for ranking input data needs (see section VI.A). This index takes into account the sensitivity of the model to the input data (e.g., deviation in ozone levels above 12 pphm, deviation in maximum ozone level), the corresponding cost for input data acquisition, and the uncertainty in the model input variables affected by the input data considered. It is important to note that this sensitivity/uncertainty index, which was used for ranking data needs in our study of the Los Angeles basin, is generally applicable; it can be used to define data needs for any urban area once the cost and uncertainty estimates have been derived. Therefore, this index constitutes a basis for a general approach to evaluate the sensitivity of the urban airshed model to the amount (and thus the quantity) of the input data. The development of a data base necessary for performing urban area model simulations would advantageously involve the information provided by the sensitivity-uncertainty analysis. Specific results obtained for the Los Angeles basin are summarized next.

B. SPECIFIC RESULTS OF THIS STUDY

The results of the sensitivity analyses were used, along with cost estimates for input data acquisition and uncertainty estimates, to define a ranking of data needs (section VI.A). The sensitivity results depend[*] on the specific simulation conditions and on the choice of the cost and uncertainty estimates (see tables VI-1 and VI-2); a choice of different sensitivity measures or cost and uncertainty estimates could lead to a different ranking of the input data needs, since the uncertainty bounds shown in figures VI-1 through VI-4 display some overlap in the estimated values of the sensitivity/uncertainty index. On the other hand, the results of a sensitivity analysis are specific to the model considered, and different factors (such as the magnitude of the perturbation, the urban area considered, the submodels--e.g., wind model--and the length of simulation time) have to be taken into account when generalizing the sensitivity results. The sensitivity results presented below should not, therefore, be interpreted as an absolute ranking of data needs; rather they (1) exemplify the general methodology developed, and (2) provide results for the specific case studied.

The main results of the sensitivity/uncertainty analysis of the simulations of 26 June 1974 and 4 August 1975 for the Los Angeles basin are listed in table VII-1. The relative importance of upper air and surface meteorology data depends strongly upon meteorological conditions. Upper air meteorological data are ranked high for the 26 June 1974 simulation, but are only of slight importance for the 4 August 1975

---

[*] That is, on the sensitivity measure (e.g., deviation in ozone levels above 12 pphm, deviation in maximum ozone level).

TABLE VII-1. RESULTS OF THE SENSITIVITY-UNCERTAINTY ANALYSIS FOR THE LOS ANGELES BASIN (SENSITIVITY OF OZONE LEVELS ABOVE 12 pphm)

(a) Simulation of 26 June 1974

Input Data	Sensitivity/Uncertainty Index Mean Value (times 10^6)
Mobile Source--Updating Inventory	23.0
Upper Air Meteorology	21.0
Initial Conditions for RHC	8.7
Area Source	4.5
Upper Air Quality	4.1
HC Speciation	2.9
Point Source	2.6
Surface Meteorology	1.3
Surface Air Quality	0.8
Mobile Source--Total Inventory	0.8
Area Sources--Temporal Resolution	0.6

(b) Simulation of 4 August 1975

Input Data	Sensitivity/Uncertainty Index Mean Value (times 10^6)
Inital Conditions for RHC	11.7
Area Sources--Spatial Resolution	11.5
Surface Meteorology	11.0
HC Speciation	9.0
Upper Air Meteorology	4.1
Mobile Source--Total Invenotry	1.8

simulation. This can be related to the fact that the wind field is strongly affected by upper air data for the 26 June 1974 case and not much perturbed for the latter case. On the other hand, surface meteorological data are more important for the 4 August 1975 simulation than for the 26 June 1974 simulation.

The ranking of the other data sets does not vary with the simulation day. The input data sets decrease in importance in the following order (meteorological data are not included; see table VII-1 for this simulation day ranking):

> Mobile sources - updating inventory

> Initial conditions for hydrocarbons

> Area sources - spatial resolution

> Upper air quality

> Hydrocarbon speciation

> Point source emissions

> Surface air quality

> Mobile sources - total inventory

> Area source - temporal resolution

A complete description, interpretation, and evaluation of the possible limitations of these sensitivity results is presented in chapters III, V, and VI, respectively. The sensitivity results depend on the magnitude of the perturbation, unless the model sensitivity is linear. As pointed out in section VI-A, the type of sensitivity analysis considered in this study is "perturbation specific"; the results depend on the perturbation considered, and caution is advised if the results are to be applied in a different context (i.e., with a different perturbation in the input data). For instance, in sensitivity case 5 the number of air quality monitoring stations was reduced from 23 to 10 and the sensitivity results would probably be quite different had the number of stations been reduced, say, from 10 to 5, since the model sensitivity is not expected to be linear in this case.

Specific attributes of the modeling conditions also affect sensitivity analysis. In this study, these conditions are characterized by the features of the Los Angeles basin, by the submodels considered (e.g., wind model, emission inventories), and by the length of simulation time. Further sensitivity studies could be carried out to investigate the effect of these attributes on the model sensitivity; possible studies are outlined in the next section.

C. FUTURE NEEDS

Because of the complex nature of photochemical grid models, a detailed sensitivity analysis of such a model is an enormous task. Much information has been obtained in this study about the sensitivity of the airshed model predictions to the amount (and thus the quality) of input data. Further work could be carried out to improve the knowledge base regarding input data needs:

> Multiple-day simulation: As the simulation time increases, the model approaches the more realistic case where emissions provide all the precursor mass and, in the limit, the model predictions become insensitive to the initial conditions. Therefore, in a multiple-day simulation, the influence of the initial conditions on the model predictions will decrease as the simulation time increases, and the importance of the emission levels (as well as meteorological description) will accordingly increase. We could therefore expect the ranking of air quality and emission data to be notably modified with a multiple-day simulation. Useful information would be obtained by carrying out multiple-day simulation studies.

> Sensitivity studies for other urban airsheds: Specific attributes of the Los Angeles basin (meteorology, aerometric network, emissions, topography) may limit the generality of some of the results. For instance, the wind model used in the Los Angeles basin was developed for complex terrain applications and is sensitive to meteorological data at the boundaries of the airshed. In an urban area with flat terrain, another type of wind model that would show a different sensitivity to the input data could be used (e.g., interpolated model). The choice of the urban area could be made according to its specific attributes and the intended applications of the models.

> Sensitivity studies for various perturbation levels: The sensitivity results presented in this study are "perturbation specific"; since the sensitivity of the model to input data perturbations is usually nonlinear, the results

> depend on the magnitude of the perturbation. It would be of interest to investigate for some specific cases the nonlinear nature of the model sensitivity. For instance, the number of available air quality monitoring stations was reduced in the sensitivity simulation focusing on air quality data. Further sensitivity studies could provide additional information on the minimum number of air quality stations needed to obtain sufficient air quality data for the definition of adequate initial conditions and boundary conditions.

Sensitivity studies involving multiple-day simulations and other urban areas would provide useful additional information on the sensitivity of photochemical models to the richness of the information data base.

REFERENCES

ABAG (1977), "Candidate Control Measures," Air Quality Maintenance Plan, Technical Memo 5, Association of Bay Area Governments, Berkeley, California.

Anderson, G. E., et al. (1977), "Air Quality in the Denver Metropolitan Region: 1974-2000." EF77-222, Systems Applications, Incorporated, San Rafael, California.

Attaway, L. D., et al. (1976), "Maintenance Shutdown of Tail Gas Treating Unit: An Assessment of Potential SO_2 Concentrations and Related Health and Welfare Effects," TR-11700, Greenfield, Attaway and Tyler, Incorporated, San Rafael, California, and Systems Applications, Incorporated, San Rafael, California.

Blumenthal, D. L., W. H. White, and T. B. Smith (1978), "Anatomy of a Los Angeles Smog Episode: Pollutant Transport in the Daytime Sea Breeze Regime," Atmos. Environ., Vol. 12, pp. 893-907.

Burton, C. S., et al. (1976), "Oxidant/Ozone Ambient Measurement Methods: An Assessment and Evaluation," EF76-111R, Systems Applications, Incorporated, San Rafael, California.

Demerjian, K. L. (1976), "Photochemical Air Quality Simulation Modeling: Current Status and Future Prospects," Paper 16-1, International Conference on Photochemical Oxidant Pollution and Its Control, Environmental Protection Agency, Raleigh, North Carolina.

Dougherty, E. P., J. T. Hwang, and H. Rabitz (1979), "Further Developments and Applications of the Green's Function Method of Sensitivity Analysis in Chemical Kinetics," J. Chem. Phys., Vol. 71, pp. 1794-1808.

Duewer, W. H., et al. (1977), "NO_x Catalytic Ozone Destruction: Sensitivity to Rate Coefficients," J. Geophys. Res., Vol. 82, pp. 935-942.

Dunker, A. M. (1980), "The Response of an Atmospheric Reaction Transport Model to Changes in Input Functions," Atmos. Environ., Vol. 14, pp. 671-679.

Durbin, Paul, and T. A. Hecht (1975), "The Photochemistry of Smog Formation," internal paper, Systems Applications, Incorporated, San Rafael, California.

EPA (1972), "Compilation of Air Pollutant Emission Factors," AP-42, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Edinger, J. G. (1973), "Vertical Distribution of Photochemical Smog in Los Angeles Basin," Environ. Sci. Technol., Vol. 7, No. 3, pp. 247-252.

GRC (1974), "Air Quality Impacts of Electric Cars in Los Angeles," Appendix A, RM-1905-A, General Research Corporation, Santa Barbara, California.

Hayes, S. R. (1978), "Performance Measures for Air Quality Dispersion Models," EF78-93, Systems Applications, Incorporated, San Rafael, California.

Hecht, T. A., J. H. Seinfeld, and M. C. Dodge (1974), "Further Development of a Generalized Kinetic Mechanism for Photochemical Smog," Environ. Sci. Technol., Vol. 8, p. 327.

Heinrich, R., S. M. Rapoport, and T. A. Rapoport (1977), "Metabolic Regulation and Mathematical Models," Prog. Biophys. Mol. Biol., Vol. 32, pp. 1-82.

Hickey, H. R., W. D. Rowe, and F. S. Skinner (1971), "A Cost Model for Air Quality Monitoring Systems," J. Air Pollut. Control Assoc., Vol. 21, No. 11, pp. 689-693.

Hoel, P. G. (1962), "Introduction to Mathematical Statistics" (John Wiley and Sons, New York, New York).

Husar, R. B., et al. (1977), "Three-Dimensional Distribution of Air Pollutants in the Los Angeles Basin," J. Appl. Meteorol., Vol. 16, pp. 1087-1096.

Kieth, R. W., and B. S. Selik (1977), "California South Coast Air Basin Hourly Wind Flow Patterns," South Coast Air Quality Management District, El Monte, California.

Koda, M., A. H. Dogru, and J. H. Seinfeld (1979), "Sensitivity Analysis of Partial Differential Equations with Applications to Reaction and Diffusion Processes," J. Comp. Phys., Vol. 30, pp. 259-282.

Lamb, R. G., and J. H. Seinfeld (1973), "Mathematical Modeling of Urban Air Pollution--General Theory," Environ. Sci. Technol., Vol. 7, pp. 253-261.

Littman, F. E. (1978), "Regional Air Pollution Study, Emission Inventory Summarization," Rockwell International, St. Louis, Missouri.

Liu, M. K., T. C. Myers, and J. L. McElroy (1979), "Numerical Modeling of Land and Sea Breeze Circulation along a Complex Coastline," Math. Comp. Simulation, Vol. 21, pp. 359-367.

Liu, M. K., D. C. Whitney, and P. M. Roth (1976), "Effects of Atmospheric Parameters on the Concentration of Photochemical Air Pollutants," J. Appl. Meteorol., Vol. 15, ppp. 829-835.

Liu, M. K., et al. (1976), "Continued Research in Mesoscale Air Pollution Simulation Modeling: Volume I--Analysis of Model Validity and Sensitivity and Assessment of Prior Evaluation Studies," EPA-600/4-76-016a, Systems Applications, Incorporated, San Rafael, California.

Lyons, W. A., and H. S. Cole (1976), "Photochemical Oxidant Transport Mesoscale Lake Breeze and Synoptic Scale Aspects," J. Appl. Meteorol., Vol. 15, pp. 733-743.

MacCracken, M. C., and G. D. Sauter, eds. (1975), "Development of an Air Pollution Model for the San Francisco Bay Area," University of California, Livermore, California.

Mayersohn, H., et al. (1976), "Atmospheric Hydrocarbon Concentrations June-September, 1975," California Air Resources Board, Atmospheric Studies Section, Division of Technical Services, Sacramento, California.

Mayersohn, H., et al. (1975), "Atmospheric Hydrocarbon Concentrations June-September, 1974," California Air Resources Board, Atmospheric Studies Section, Division of Technical Services, Sacramento, California.

McRae, G. J., and J. W. Tilden (1980), "A Sensitivity and Uncertainty Analysis of Urban Scale Air Pollution Models--Preliminary Steps," Proc. Second Joint Conference on Applications of Air Pollution Meteorology, AMS/APCA, 24-27 March 1980, New Orleans, Louisiana.

Miedema, A. K., et al. (1973), "Cost of Monitoring Air Quality in the United States," EPA-450/3-74-029, Research Triangle Institute, Research Triangle Park, North Carolina.

Reynolds, S. D., and P. M. Roth (1980), "The Systems Applications, Incorporated Airshed Model--A Review and Assessment of Recent Evaluation and Applications Experience," submitted to J. Air Pollut. Control Assoc.

Reynolds, S. D., P. M. Roth, and J. H. Seinfeld (1973), Atmos Environ., Vol. 7, p. 1033.

Reynolds, S. D., et al. (1979), "Photochemical Modeling of Transportation Control Strategies, Vol. 1, Model Development, Performance Evaluation and Strategy Assessment," EF79-37, Systems Applications, Incorporated, San Rafael, California.

Reynolds, S. D., et al. (1976), "Continued Development and Validation of a Second Generation Photochemical Air Quality Simulation Model: Volume II--Refinements in the Treatment of Chemistry, Meteorology, and Numerical Integration Procedures," EF75-24R, Systems Applications, Incorporated, San Rafael, California.

Russell, P. B., and E. E. Uthe (1978), "Regional Patterns of Mixing Depth and Stability: Sodar Network Measurements for Input to Air Quality Models," Bull. Am. Meteorol. Soc., Vol. 59, pp. 1275-1287.

SCAPCD (1976), "Fuel Use and Emissions from Stationary Combustion Sources," Southern California Air Pollution Control District, El Monte, California.

Seigneur, C. (1978), "Mathematical Analysis of Air Pollution Models," Ph.D. Thesis, University of Minnesota, Minneapolis, Minnesota.

Skelton, E. P., et al. (1977), "Methodology for Estimating Emissions for the Sacramento AQMP," Sacramento Air Pollution Control District, Sacramento, California.

Souten, D. R., G. E. Anderson, and R. G. Ireson (1980), "National Commission on Air Quality Los Angeles Regional Study SAI Task 1 Report: Review of the SIP Process and Framework (1.A.) Review and Analysis of South Coast Air Quality Management District Air Quality Management Plan Modeling Methods (1.B)," Draft Final Report, EF80-17, Systems Applications, Incorporated, San Rafael, California.

Tesche, T. W. (1978), "Evaluating Simple Oxidant Prediction Methods Using Complex Photochemical Models," EM78-14, Systems Applications, Incorporated, San Rafael, California.

Tesche, T. W., and M. A. Yocke (1978), "Numerical Modeling of Wind Fields over Mountainous Regions in California," Conference on Sierra Nevada Meteorology, American Meteorological Society, 19-21 June 1978, South Lake Tahoe, California.

Unger, C. D. (1976), "Meteorological Input into Photochemical Models," unpublished results, California Air Resources Board, Sacramento, California.

West, R. (1979), private communication from Richard West, CSBE, to William Rogers Oliver, Systems Applications, Incorporated, San Rafael, California.

Whitten, G. Z., and H. Hogo (1977), "Mathematical Modeling of Simulated Photochemical Smog," EPA-600/3-77-011, Systems Applications, Incorporated, San Rafael, California.

Yotter, E. E. (1979), "Motor Vehicle Emissions Assessment Using Caltrans Transportation Systems Analysis (Modeling) Data," California Air Resources Board, Planning Division, Air Quality Maintenance Planning Branch, Sacramento, California.